Thesis for The Degree of PhD of Engineering

Digital Compensation Techniques for Power Amplifiers in Radio Transmitters

Jessica Chani-Cahuana

Communication and Antenna Systems Group
Department of Electrical Engineering
Chalmers University of Technology
Göteborg, Sweden, 2017

Digital Compensation Techniques for Power Amplifiers in Radio Transmitters

Jessica Chani-Cahuana

ISBN 978-91-7597-596-2
Doktorsavhandlingar vid Chalmers Tekniska Högskola
Series No: 4277
ISSN 0346-718X

Communication and Antenna Systems Group
Department of Electrical Engineering
Chalmers University of Technology
SE-412 96 Göteborg, Sweden
Phone: +46 (0) 31 772 1000

This thesis has been prepared using LaTeX
Printed by Chalmers Reproservice
Göteborg, Sweden 2017

Abstract

Power amplifiers (PAs) are vital components in radio transmitters because they are responsible to amplify the low power communication signals to power levels suitable for transmission. Important requirements of PAs are high efficiency and linearity. Unfortunately, there is a tradeoff between efficiency and linearity. In order to satisfy both requirements, designers prefer to prioritize the efficiency in the design process while the linearity is taken care of later using external linearization techniques. Among the linearization techniques proposed in the literature, digital predistortion (DPD) has drawn a large attention of the industrial and academic sectors because it can provide a good compromise between linearity, implementation complexity and efficiency. This thesis treats different aspects related to the compensation of PA nonlinear distortion through DPD.

One issue in the synthesis of DPD is that the optimal output from a predistorter is unknown. To overcome this problem, the concept of iterative learning control (ILC) for the linearization of PAs is introduced. An ILC scheme is derived that is able to identify the optimal predistorted signal that linearizes a PA. Based on the ILC framework, a novel approach to derive model structures for digital predistorters is proposed.

Techniques to identify the parameters of digital predistorters have been developed. Three parameter identification techniques based on ILC have been proposed: an offline technique that can be used for research purposes to select proper models for predistorters, an adaptive technique that is able to achieve better performance than conventional identification techniques used in DPD, and an identification technique that allows us to estimate the predistorter parameter using only one of the in-phase/quadrature (IQ) components of the PA output signal.

The issue of gain normalization in the indirect learning architecture (ILA) has been investigated. A variant to ILA that eliminates the need for a normalization gain and simplifies the DPD synthesis is proposed.

Performance limits on PA linearization has also been investigated and an expression for the lower bound for the normalized mean square error (NMSE) performance has been derived.

The improved linearity performance achieved through the techniques devel-

oped in this thesis can enable a better utilization of the potential performance of existing and emerging highly efficiency PAs, and are therefore expected to have an impact in future wireless communication systems.

Keywords: digital predistortion, power amplifier, nonlinear, efficiency, Volterra series.

List of Publications

Appended Publications

This thesis is based on the work contained in the following papers.

[A] **J. Chani-Cahuana**, P. Landin, C. Fager, and T. Eriksson, "Iterative Learning Control for the Linearization of Power Amplifiers", in *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 9, pp. 2778-2789. Sept., 2016.

[B] **J. Chani-Cahuana**, P. Landin, C. Fager, and T. Eriksson, "Structured Digital Predistorter Model Derivation using Iterative Learning Control", *European Microwave Conference*, London, 2016, pp. 178-181.

[C] **J. Chani-Cahuana**, C. Fager, and T. Eriksson, "A New Variant of the Indirect Learning Architecture for the Linearization of Power Amplifiers", *IEEE European Microwave Integrated Circuits Conference*, Paris, 2015, pp. 444-447.

[D] **J. Chani-Cahuana**, M. Özen, C. Fager, and T. Eriksson, "Digital predistortion parameter identification technique using real-valued measurement output data", in *IEEE Transactions on Circuits and Systems II: Express Briefs*, March, 2017.

[E] **J. Chani-Cahuana**, P. Landin, C. Fager, and T. Eriksson, "On the Behavior of the Normalized Mean Square Error in Power Amplifier Linearization", *Submitted to IEEE Microwave and Wireless Components Letters.*

Other Publications

The following papers have been published but are not included in the thesis. The content partially overlaps with the appended papers or is out of the scope of this thesis.

[a] D. Gustafsson, **J. Chani-Cahuana**, D. Kuylenstierna, I. Angelov, N. Rorsman, C. Fager, "A Wideband and Compact GaN MMIC Doherty Amplifier for Microwave Link Applications," in *IEEE Transactions on Microwave Theory and Techniques* , vol. 61, no. 2, pp. 922-930, Feb., 2013

[b] D. Gustafsson, **J. Chani-Cahuana**, D. Kuylenstierna, I. Angelov, and C. Fager, "A GaN MMIC Modified Doherty PA with Large Bandwidth and Reconfigurable Efficiency," *IEEE Transactions on Microwave Theory and Techniques* , vol. 61, no. 2, pp. 922-930, Feb., 2013

[c] C. M. Andersson, D.Gustafsson, **J. Chani-Cahuana**, R. Hellberg, and C. Fager, "A 1-3 GHz Digitally Controlled Dual-RF Input Power Amplifier Design Based on a Doherty-Outphasing Continuum Analysis," *IEEE Transactions on Microwave Theory and Techniques* , vol. 61, no. 10, pp. 3743-3752, Oct., 2013

[d] **J. Chani-Cahuana**, P. Landin, D. Gustafsson, C. Fager, and T. Eriksson, "Linearization of Dual-Input Doherty Power Amplifiers", *IEEE International Workshop on Integrated Nonlinear Microwave and Milimetrewave Circuits (INMMiC)*, Leuven, April, pp. 1-3.

[e] T. Eriksson, C. Fager, P. N. Landin, U. Gustavsson, **J. Chani-Cahuana** and K. Hausmair, "Linearization of Difficult Cases - MIMO, GaN and Deep Compression", *Gigahertz Symposium*, Gothenburg, Sweden, January, 2014

[f] C. Fager, X. Bland, K. Hausmair, **J. Chani-Cahuana**, and T. Eriksson, "Prediction of Smart Antenna Transmitter Characteristics Using a New Behavioral Modeling Approach", *IEEE MTT-S International Microwave Simposium*, Tampa, FL, 2014, pp. 1-4.

[g] A. Soltani Tehrani, **J. Chani**, T. Eriksson, and C. Fager, "Investigation of Parameter Adaptation in RF Power Amplifier Behavioral Models", *arXiv:1410.8127v1*, October, 2014

[h] M. Pampin-Gonzalez, M. Özen, C. Sanchez-Perez, **J. Chani-Cahuana**, and C. Fager, "Outphasing Combiner Synthesis from Transistor Load Pull Data", *IEEE MTT-S International Microwave Symposium*, Phoenix, AZ, 2015, pp. 1-4.

Theses

[i] **J. Chani-Cahuana**, C. Fager, and T. Eriksson, "Digital Predistortion for the Linearization of Power Amplifiers", *Chalmers University of Technology*, Thesis for the licentiate degree, Gothenburg, 2015

As part of the author's doctoral studies, some of the work presented in this thesis has been previously published in [i]. Figures, tables and text in [i] might therefore be fully or partly reproduced in this thesis.

Abbreviations

Abbreviations

1G	First Generation
4G	Fourth Generation
5G	Fifth Generation
ACPR	Adjacent Channel Power Ratio
ADC	Analog-to-Digital Converter
DAC	Digital-to-Analog Converter
DC	Direct Current
DLA	Direct Learning Architecture
DPD	Digital Predistortion
EEMP	Extended Envelope Memory Polynomial
EMP	Envelope Memory Polynomial
EVM	Error Vector Magnitude
GMP	Generalized Memory Polynomial
ILA	Indirect Learning Architecture
ILC	Iterative Learning Control
ILC-DPD	Iterative Learning Control based Digital Predistortion
IQ	In-phase/Quadrature
LS	Least Squares
LTE-A	Long Term Evolution Advanced
MILA	Model-based Indirect Learning Architecture
MP	Memory Polynomial
NMSE	Normalized Mean Square Error
PA	Power Amplifier
PAPR	Peak-to-Average Power Ratio
RBS	Radio Base Station
RF	Radio Frequency
RILC-DPD	Real-valued Iterative Learning Control based DPD
SNR	Signal-to-noise ratio
V-DDR	Volterra-Dynamic Deviation Reduction
VS-GMP	Vector Switched Generalized Memory Polynomial

Contents

Part I

Overview

Chapter 1

Introduction

In the last decades, we have witnessed the evolution of wireless communication systems and experienced how the introduction of those technologies has progressively changed the way we communicate and access information. Good examples of this are mobile systems which have taken us from the first-generation (1G) analogue communication systems which could only handle voice calls, all the way to the sophisticated services offered by the fourth generation (4G) long-term evolution advanced (LTE-A) systems which include high-speed mobile internet, mobile video streaming, mobile cloud-based applications, etc. Now, we are on the way to fifth-generation (5G) systems which promises to go beyond connectivity, providing wireless connection to any application, service or device anywhere at any time [1, 2]. But the fast growing demands for wireless services, besides of becoming a core part of most people's lives has dramatically increased the energy consumption of mobile networks. Reports indicated that in 2016, the largest mobile network in the world consumed over 19.7 billion KWh of energy [3, 4] which is almost the same amount of power that the Island of Taiwan spends in a month and also equates to the emission of 14 million tons of carbon dioxide [5, 6]. These alarming results of environmental pollution together with the interest of mobile network operators to reduce their electricity expenses has lead to an increased attention to the energy efficiency within wireless networks [7, 8].

Major contributors to the energy consumption in wireless networks are the radio base stations (RBS) which account for 80 % of the total energy consumption across the network [7], a large amount of which is wasted due to the inefficient operation of the radio equipment and power amplifiers (PAs) [9]. Improving the power efficiency of PAs plays an important role on the reduction of the energy consumption of RBSs. Over the years researchers have studied and proposed different ways to achieve that goal, but different problems associated with the operation of PAs and contradictions with other system requirements have made the work difficult [10, 11]. But before explaining the rationale behind this, it is important to understand what PAs are and what role they play

in RBSs.

Power amplifier (PAs) are important components of RBSs because they are responsible to increase the power of the communication signals to power levels that are suitable for transmission. Being one of the last components in the radio transmitter chain, PAs handle the major amount of power in the entire chain [12]. For this reason, the efficiency at which they convert direct current (DC) power into radio frequency (RF) power plays an important role on the overall power consumption of RBSs. Another important fact about PAs is that besides producing RF power they emit a large amount of heat due to their inefficient operation. In order to avoid overheating of the other components in the system, PAs generally require an air-conditioning unit which increases even more the energy consumption in RBSs [9].

Another important requirement of PAs is the linearity of their output response [12]. Linearity at the output of PAs is important for two reasons: to fulfill the bit error rate requirements of the system and to satisfy the stringent spectral requirements which limit the amount of distortion that may be leaked into neighboring channels [13]. The latter is of paramount importance due to the scarcity of the frequency spectrum and the commercial competition between network operators who pay millions for the exclusive use of a small portion of the spectrum [10]. Unfortunately, linearity and efficiency are two conflicting requirements. This is because in order to improve their efficiency, conventional PAs must be operated close to saturation where they present a strong nonlinear behavior [10,14]. The nonlinear behavior of PAs not only distorts the transmitted signal but also generates spectral regrowth which causes interference to signals transmitted in neighboring channels, as can be seen in Fig. 1.1. In order to improve the linearity, PAs must be backed-off far from their saturation point where they operate with low power efficiency. This situation combined with the large peak-to-average power ratio (PAPR) presented by modern, spectral-efficient communication signals results in very low overall efficiency numbers [15].

This linearity-efficiency tradeoff is so critical that in order to meet both requirements, system designers prefer to operate PAs at high-efficiency levels distorting the peaks of the communication signal and remove later the distortion [14]. This convention gives rise to two major areas in the research related to PAs. The first one is driven by the need to develop high-efficiency PA architectures that comply with the operating frequencies, bandwidth and output power requirements of modern wireless systems, while the second is devoted to develop techniques to compensate for the distortion that the high-efficiency PAs introduce. It is in the latter where the work presented in this thesis takes place.

Over the years different methods have been developed to compensate the nonlinear distortion introduced by PAs [11]. Among these methods, digital predistortion (DPD) has drawn the most attention in recent years partly due to the new possibilities opened up by the advancements of high speed digital

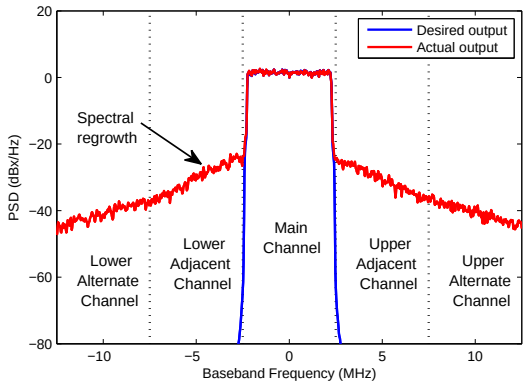


Figure 1.1. Spectra of the output signal of a Class AB power amplifier driven with a 5 MHz-LTE signal. Note that due to the nonlinearity of the PA, spectral regrowth is generated in the neighboring channels.

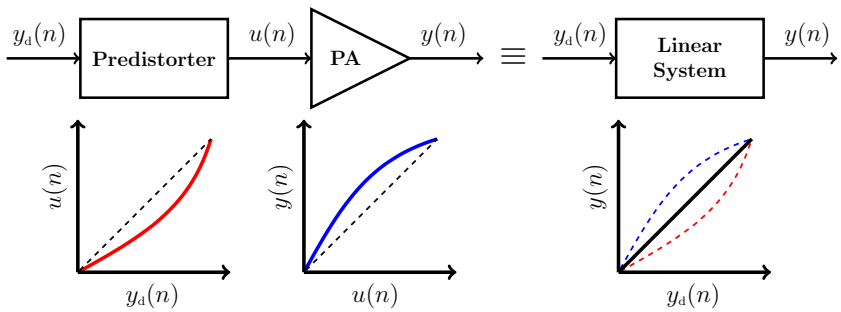


Figure 1.2. Operation principle of digital predistortion

signal processing technologies. The idea of DPD is to compensate the nonlinear behavior by distorting the amplitude and phase characteristics of the communication in such a way that when the predistorted signal is sent to the PA, the output response results in a linear amplification of the signal to be transmitted. In DPD, this is done by introducing, before the PA, a digital nonlinear block which contains the PA inverse characteristics, as is depicted in Fig. 1.2. This concept, although simple, has proven to be effective providing a good compromise between linearity, efficiency and implementation complexity [11, 14, 16] and will be the main focus of this work.

1.1 Thesis contribution

The thesis makes five distinct contributions to the field of DPD.

An issue encountered in DPD is that the optimal output from the predis-

torter is unknown beforehand [17]. To alleviate this problem, in Paper [A], we introduce the concept of iterative learning control (ILC) to the linearization of PAs and propose a new parameter identification technique based on ILC which focuses on finding the optimal predistorted signal that linearizes the PA before estimating the predistorter parameters. Based on experimental results, it was shown that for the most difficult linearization cases, the proposed ILC scheme can successfully identify the optimal signal that linearizes a PA. It was also shown that the proposed ILC-based parameter identification technique can provide better linearity performance than existing techniques when the PA are in deep compression and when the output signal has low signal-to-noise ratio (SNR).

An important step in the design of a digital predistorter is the selection of the model used in the predistorter. In Paper [B], a novel structured technique to analytically derive inverse model structures for PAs is proposed. By using the ILC concept, the proposed technique first derives an analytical expression of the predistorted signal and uses it to derive, in a structure manner, predistorter models from Volterra-based PA models. Experimental results showed that this technique can derive models that provide better linearity performance than conventional models used in DPD.

A critical issue encountered in the synthesis of predistorters based on the indirect learning architecture (ILA) has always been the selection of the normalization gain. Different ways to compute that gain have been proposed [18–21], but in general there is not a clear consensus on how to select it or on how the selection affects the linearization performance. In Paper [C], the effects of the normalization gain are investigated and a variant to the ILA is proposed that eliminates the need of the normalization gain while allowing improved control of the PA output power.

The adoption of wider transmission bandwidths creates new challenges in the implementation of DPD solutions [22]. In wideband DPD systems, more expensive and faster analog-to-digital converters (ADCs) are required to effectively linearize the PAs. To reduce the requirements on ADCs, different solutions based on undersampling [23–26] and band-limited modeling [27, 28] have been proposed. In Paper [D], we present a completely different approach to this problem. There we propose a novel parameter identification technique that requires only one of the in-phase/quadrature (IQ) components of the PA output signal. Since only one of the IQ components needs to be acquired, the suggested technique may help to reduce the number of ADCs required in DPD feedback receivers.

In DPD, the performance assessment is generally done by comparing the performance of a proposed technique to existing solutions without taking into account how far the performance is from the ideal one to see if further improvements is necessary. In Paper [E], we derived an analytical expression for the lower bound for the normalized mean square error (NMSE) obtained in linearized PAs. The derived expression gives us a better insight into the

behavior of the NMSE with respect to the output power from the PA and provides a reference with which to compare the performance of DPD linearization schemes.

1.2 Thesis outline

The rest of the thesis is organized as follows. In Chapter 2, a brief introduction to the behavioral modeling of PAs is presented. In Chapter 3, basic concepts of DPD are presented. The problem of DPD is mathematically formulated. A short description of a DPD system is presented and commonly used linearity metrics are introduced. Finally, a novel PA linearization technique based on ILC is presented.

In Chapter 4, model structures used in DPD are discussed and a new approach to the design of predistorter model structures is presented.

In Chapter 5, different techniques to identify the parameters of digital predistorters are discussed. After reviewing conventional identification techniques such as ILA and the direct learning architecture (DLA), three novel parameter identification techniques based on the ILC framework are presented. The first one is an offline technique that can be used to select models for digital predistorters. The second one extends the first technique for real-time DPD scenarios. The third technique is an identification technique which allows us to estimate the predistorter parameters using only one of the IQ components of the PA output signal. The gain normalization issue in ILA is also discussed in this chapter.

Chapter 6 deals with performance limits in PA linearization. A closed-form expression for the NMSE performance is presented. Finally, in Chapter 7, conclusions from the research done are drawn, the contributions of the appended papers are presented and future work in the field is discussed.

Chapter 2

Power Amplifier Behavioral Modeling

Digital predistortion and PA behavioral modeling are two research areas that are closely related. This is because in order to compensate the distortion introduced by PAs, it is important to find an accurate way to characterize their nonlinear behavior, here referred to as PA forward behavior. A behavioral model, also known as empirical model or black-box model, is a model that characterizes the behavior of a system relying only on a set of input-output observations [29]. In this chapter, a short introduction to the behavioral modeling of PAs is presented. Note that it is not the author's intention to provide a full survey on this topic, for more information the reader is referred to [29–31] and the references within.

2.1 Power amplifier nonlinear behavior

PAs are important components in the radio transmitter chain because they are responsible to amplify the communication signals to power levels suitable for transmission. Ideally, it is desired that the amplification is done so that the output is a linearly scaled version of the PA input signal, in reality, those devices present a nonlinear behavior which is more accentuated as they are operated closer to their saturation point. Fig. 2.1 shows a typical input and output amplitude characteristic of a PA.

The nonlinear behavior of PAs has two major components, a static nonlinearity and dynamic distortions [32]. The static nonlinearity is the major source of distortion in PAs which is shown as the compression of the output signal amplitude as the input amplitude is increased. This nonlinear compressing behavior is mainly attributed to the nonlinear DC characteristics of the active device or transistor [30]. The dynamic distortions also known as memory effects are less dominant but when they are present, the output signal

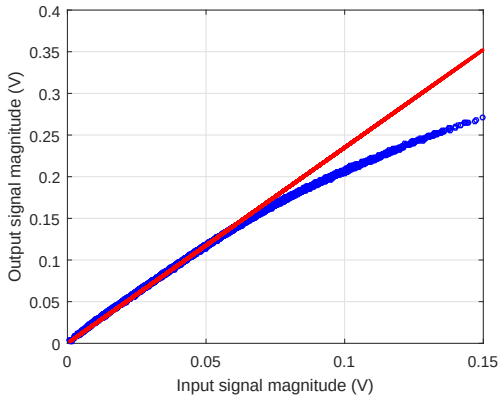


Figure 2.1. Input and output amplitude characteristics of a practical power amplifier.

do not only depend on the current input sample but also on previous input samples [32]. The memory effects are attributed to different sources, e.g., the frequency response of the matching networks and device parasitics, trapping effects, temperatures changes due to the power dissipation in the active device, to mention a few [30, 32].

Although the contributions of the static nonlinearity are more dominant than the memory effects, they are equally important especially in DPD where both need to be compensated for to be able to achieve acceptable levels of distortion [32].

2.2 Power amplifier model structures

Over the years several PA behavioral models have been proposed in the literature, going from simple models that can only characterize the static nonlinearity of PAs [33–35], to more elaborate models that can also account for memory effects, such as the Volterra series [36], different reduced forms of the Volterra series [16, 17, 31, 37], neural networks [30], to mention a few. In this work, the main focus is on models based on the Volterra series.

Note that all models considered in this work are constructed using the complex-valued baseband equivalent signals of the RF PA input and output signals. Although PAs used in wireless communication systems are nonlinear functions that map a real-valued RF (or bandpass) signal to a real-valued RF output, because the PA input signal is narrowband in relation to the RF carrier and only the information appearing close to the RF carrier is of relevance, the behavior of PAs can be translated into the baseband domain [38]

The relation between a narrowband bandpass signal $u_{\mathrm{RF}}(t)$ and its complex-

valued baseband equivalent $u(t)$ is given by

$$\begin{aligned} u_{\mathrm{RF}}(t) &= A(t)\cos\big(\omega_c t + \phi(t)\big) \\ &= Re\big\{A(t)e^{j(\omega_c t+\phi(t))}\big\} \\ &= Re\big\{u(t)e^{j\omega_c t}\big\} \end{aligned} \tag{2.1}$$

with $u(t) = A(t)e^{j\phi(t)}$. $A(t)$ and $\phi(t)$ denote the amplitude and phase modulation, and ω_c denotes the RF carrier angular frequency [39].

The Volterra series

When talking about PA behavioral modeling, probably the first model that comes to mind is the Volterra series [36]. This is because it is one of the first models considered to characterize the PA with memory effects and is the foundation for other PA behavioral models in the literature. The Volterra series is a mathematical tool used to describe the behavior of time-invariant nonlinear dynamic systems with fading memory [36]. It is considered to be the extension of the impulse response concept from linear systems to nonlinear systems. The discrete time complex-baseband Volterra series can be formulated as

$$\begin{aligned} y(n) = \sum_{\substack{p=1 \\ p\ \text{odd}}}^{P} \sum_{m_1=0}^{M} \sum_{m_2=m1}^{M} \cdots \sum_{m_{(p+1)/2}=m_{(p-1)/2}}^{M} \\ \sum_{(p+3)/2=0}^{M} \cdots \sum_{m_p=m_{p-1}}^{M} h_p(m_1, m_2, \ldots, m_p) \\ \prod_{i=1}^{(p+1)/2} u(n-m_i) \prod_{j=(p+3)/2}^{p} u^*(n-m_j) \end{aligned} \tag{2.2}$$

where $h_p(m_1, \cdots, m_p)$ are the parameters of the Volterra series, more formally known as Volterra kernels. $(\cdot)^*$ represents the complex conjugate, P is the nonlinear order and M is the memory depth. This model has the advantage of being linear in the parameters.

The Volterra series can provide good model accuracy but, as noticed from (2.2), its number of parameters increases drastically with the nonlinear order P and memory depth M, which limits its application to weakly nonlinear PAs. In order to reduce the computational complexity, several models have been developed to simplify the structure of the Volterra series. These models will be treated in the following section.

Pruned-Volterra series models

Pruned-Volterra models also known as reduced-Volterra models are model structures that contain a subset of the basis functions of the Volterra series.

Different approaches have been proposed to prune the terms of the Volterra series. In early works, the pruning was done more or less in adhoc manner by choosing structures that provided reasonable accuracy [40], later works then incorporating physical knowledge of PAs to prune the Volterra series in a more structured way [38, 41, 42]. The literature on pruned-Volterra models is extensive, in this section we will focus on some of the most widely-known models which are also used throughout this work.

The most commonly known reduced Volterra model is probably the *memory polynomial* (MP) model. Proposed in [17], the MP model is a reduction of the Volterra series in which only products with the same time-shifts are included [32]. The MP model can be formulated as

$$y(n) = \sum_{p=1}^{P} \sum_{m=0}^{M} a_{pm} x(n-m)|x(n-m)|^{p-1} \tag{2.3}$$

where a_{pm} are the model parameters. $|\cdot|$ denotes the absolute value. P and M represent the maximum nonlinear order and the memory depth of the model, respectively.

Another important model in this category is the *generalized memory polynomial* (GMP) [16]. This model extends the MP model by also introducing products with different time-shifts, which are generally referred to as cross terms. The GMP model can be written as [43]

$$\begin{aligned} y(n) = &\sum_{p=1}^{P} \sum_{m=0}^{M-1} a_{pm} u(n-m)|u(n-m)|^{p-1} \\ &+ \sum_{p=2}^{P} \sum_{m=0}^{M-1} \sum_{\substack{l=\max\{-m,-L\} \\ g\neq 0}}^{L} b_{pml} u(n-m)|u(n-m-l)|^{p-1} \end{aligned} \tag{2.4}$$

where a_{pm} and b_{pml} are the model parameters. P, M, and L are the nonlinear order, memory length and cross-term length, respectively. Similar to the Volterra series, the MP and GMP are also linear in the parameters, which means that their parameters can be estimated using least squares techniques.

Other models in this category include the Volterra dynamic deviation reduction (V-DDR), the envelope memory polynomial (EMP) [44], the extended EMP model (EEMP) [42], the Wiener and Hammmerstein models [16, 32] to mention a few.

2.3 Model parameter estimation

Once a behavioral model for the PA is chosen, the next step is the estimation of its parameters. The estimation technique used depends on the structure of the model. For models that are linear in the parameters, such as the Volterra

series and most of the reduced Volterra models, the linear least square (LS) estimator is generally used [30].

The LS approach estimates the parameters in order to minimize the sum-squared error between the observed data $y(n)$ and the model output $\hat{y}(n)$, i.e.

$$J(\theta) = \sum_{n=0}^{N-1} e(n)^2 = \sum_{n=0}^{N-1} \left| y(n) - \hat{y}(n) \right|^2 \tag{2.5}$$

where N is the number of samples of the input $u(n)$ and output $y(n)$ signals.

Models that are linear in the parameters can be written more compactly as

$$\hat{\mathbf{y}} = \mathbf{H}\boldsymbol{\theta} \tag{2.6}$$

where $\mathbf{y}$ is a column vector containing the samples of the model output $\hat{y}(n)$, $\mathbf{H}$ is a matrix consisting of the basis functions of the model, and $\boldsymbol{\theta}$ is a column vector containing the model parameters.

The LS solution is readily given by [45]

$$\hat{\boldsymbol{\theta}} = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{y} \tag{2.7}$$

where $\mathbf{y}$ is a vector containing the samples of the observed output signal $y(n)$ and $(.)^H$ denotes the Hermitian transpose.

Chapter 3

Digital predistortion

DPD is currently the most active research area for the linearization of PAs because it offers a good tradeoff between implementation complexity and performance. By introducing a nonlinear block that contains the inverse behavior of the PA, DPD is able to compensate the nonlinear distortion generated by a PA. This chapter is thought to provide an introduction to the DPD of PAs. This chapter starts with the formulation of the DPD problem. Next, a short description of a DPD system is presented. Thereafter, performance metrics that are commonly used to evaluate the linearity of PAs are reviewed. Finally, a novel PA linearization scheme based on ILC is introduced.

3.1 Formulation of the digital predistortion problem

Broadly speaking, there are two approaches to synthesize a digital predistorter: to find and analytically invert a forward model of the PA, or to select a model structure to realize the predistorter function and estimate its parameters using some kind of identification technique [32, 38].

The first approach was considered in early DPD studies, when the Volterra series was used as PA behavioral model [46,47]. Then, the inverse of a Volterra model of the PA was computed using the p-th order inverse theory [48], which is a computationally heavy technique to invert the nonlinearity of a Volterra model up to the p-th nonlinear order. Due to the complexity of computing an analytical inverse of a PA forward model, and the introduction of parameter identification techniques such as the indirect learning architecture (ILA) [47] which were more simple to implement, the first approach was largely left behind and the second became more or less the norm in the synthesis of digital predistorters.

Based on that, the problem of DPD can be graphically represented as shown in Fig 3.1. Consider a PA system defined by $y(n) = F_{\mathrm{PA}}[u(n)]$. For this

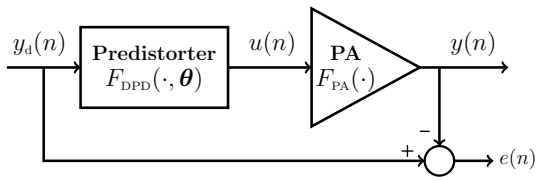


Figure 3.1. Optimisation problem of digital predistorter

system, the goal is to find a predistorter function denoted by $F_{\mathrm{PD}}[y_{\mathrm{d}}(n), \boldsymbol{\theta}]$, so that the output $y(n)$ from the cascade of the predistorter and PA system is as close as possible to a desired output response $y_{\mathrm{d}}(n)$, where close is measured in the sense of a suitable norm. This can be formulated as an optimisation problem as follows

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \big\| e(n) \big\| = \arg\min_{\boldsymbol{\theta}} \big\| y_{\mathrm{d}}(n) - F_{\mathrm{PA}}\left[F_{\mathrm{PD}}[y_{\mathrm{d}}(n), \boldsymbol{\theta}] \right] \big\| \tag{3.1}$$

In the next two chapters, we will discuss the steps taken to synthesize digital predistorters using the second approach. Chapter 4 is dedicated to model structures used in DPD, and Chapter 5 will treat the techniques used to identify the parameters of predistorter models.

3.2 DPD system description

In DPD studies, DPD systems are depicted as a simple cascade of a predistorter and a PA, as the one shown in Fig. 3.1. This section provides a description of a practical DPD system in more detail and also describes the measurement setup used in our experiments.

A block diagram of a DPD system is depicted in Fig. 3.2 [49]. The signal to be transmitted is passed through the predistorter generating the digital baseband predistorted signal. This signal is converted to the analog domain using a pair of digital to analog converters (DACs) to then be up-converted to the RF carrier frequency using an IQ modulator. Thereafter, the RF bandpass signal is sent to a pre-driver which amplifies the signal to power levels suitable to drive the PA. In order to synthesize the predistorter, a portion of the PA output signal is extracted, down-converted and digitized. The measurement circuitry used for this purpose is referred to as DPD feedback receiver. The resulting digital baseband signal is then sent to a parameter identification block which estimates the predistorter parameters and updates them to the predistorter.

Because the PA nonlinear behavior causes bandwidth expansion of the PA output signal, the DPD feedback receiver must cover a span that is a multiple of the communication signal bandwidth equivalent to the nonlinearity order to be compensated for. Typically, a bandwidth five times wider is used [32].

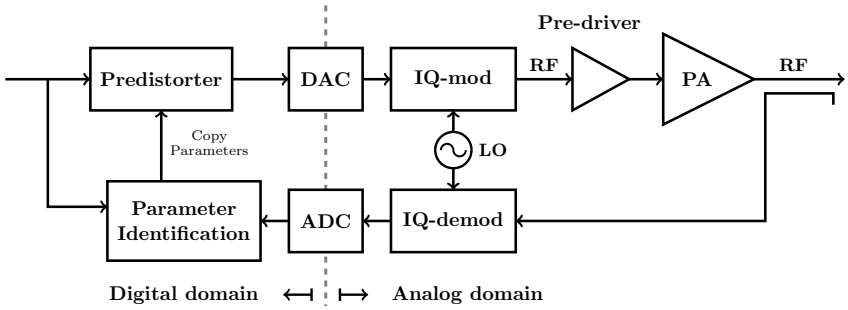


Figure 3.2. Block diagram of a digital predistortion system

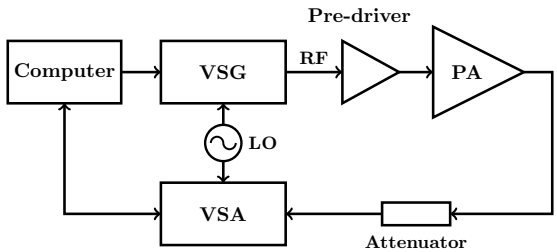


Figure 3.3. Block diagram of typical measurement setup for digital predistortion

In our experiments, the DPD system is emulated using the measurement setup shown in Fig. 3.3. The digital predistorter and all the signal processing involved in the synthesis of a predistorter are implemented in a computer using MATLAB. The baseband predistorted signal is downloaded into a vector signal generator which sends the RF modulated signal to the pre-driver and PA chain. The PA output signal is acquired using a signal analyzer which sends the baseband output signal back to the computer.

In Papers [A-D], the signal generator and signal analyzer used in the measurement setups were a Keysight E4438C vector signal generator, and a Keysight N9030A PXA signal analyzer, respectively. In Paper [E], the experiments were run using RF WebLab, which is a remote-access measurement setup provided by Chalmers University of Technology and National Instruments. RF WebLab is available at www.dpdcompetition.com [50].

3.3 Linearity performance metrics

In order to be able to evaluate the linearity of PAs, it is necessary to define metrics that measure the amount of distortion PAs introduced. These performance metrics are typically defined by wireless communication standard regulations not only to maintain a suitable system performance but also to en-

sure not to interfere with wireless systems operating in neighboring channels. This section presents the most commonly used criteria in the DPD community to evaluate the linearity performance of PAs.

Normalized mean square error

The normalized mean square error (NMSE) is defined as

$$\text{NMSE} = \frac{\sum_{n=0}^{N-1} | \, y(n) - \hat{y}(n)|^2}{\sum_{n=0}^{N-1} |y(n)|^2} \tag{3.2}$$

where $y(n)$ denotes the measured signal at the PA output and $\hat{y}(n)$ denotes the modeled output. The NMSE is a full-band measure, but due to the high dynamic range of the PA stimuli, in practice it is used as an in-band measure [51].

Error vector magnitude

The error vector magnitude (EVM) is a performance metric that is widely adopted in wireless communication standards, but it is not commonly used in DPD studies. Unlike the NMSE, the EVM is a true in-band performance metric. The EVM is defined as [52]

$$\text{EVM} = \frac{\sum |Y(k) - Y_\text{d}(k)|^2}{\sum |Y_\text{d}(k)|^2} \tag{3.3}$$

where $Y_\text{d}(k)$ is the constellation points extracted from the reference signal $y_\text{d}(n)$ after demodulation and $Y(k)$ is the constellation extracted from the measured output signal $y(n)$.

Adjacent channel power ratio

The adjacent channel power ratio (ACPR) is an out-of-band performance metric. It measures the power of the distortion components that are leaked into the adjacent channel in relation to the power of the signal in the main channel [30]. The ACPR is defined as

$$\text{ACPR} = \max_{m=1,2} \left[\frac{\int_{(\text{adj})_m} |Y(f)|^2}{\int_\text{ch.} |Y(f)|^2} \right] \tag{3.4}$$

where $Y(f)$ denotes the power spectrum of the measured output signal $y(n)$. The integration in the numerator is done over the adjacent channel that presents the largest power and the integration in the denominator is done over the transmission channel.

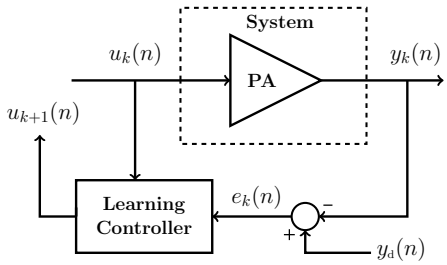


Figure 3.4. Iterative learning control scheme for the linearization of power amplifiers.

3.4 Iterative learning control scheme for PA linearization

The ultimate goal of a predistorter is to generate an optimal predistorted signal $u_{\text{opt}}(n)$ that will drive the PA, as close as possible, to a desired linear output response y_{d}. In DPD, however, the optimal predistorted signal/output of the optimal predistorter $u_{\text{opt}}(n)$ is unknown. For that reason, predistorters are designed using iterative schemes based only on the evolution of the input and output signals from the PA. To overcome this problem, in Paper [A], we proposed an iterative learning control (ILC) scheme which is able to estimate such optimal predistorted signal $u_{\text{opt}}(n)$.

ILC is a technique to iteratively estimate the optimal input signal that drives a system to a desired output response. This technique is based on the idea that the performance of a system executing the same task repeatedly can be improved by learning from previous operations [53]. If the operating conditions of a system are the same each time it is executed, any error observed in the output response will be repeated every time the system is executed. That information can then be used to modify the input signal to reduce the error obtained the next time the system is operated [54]. ILC differs from other learning type-techniques in that ILC does not modify a controller or a set of parameters of a controller, instead it directly modifies the input signal to the system [53].

The proposed ILC scheme for PA linearization is depicted in Fig. 3.4, where the subscript k denotes the iteration number. The goal of the scheme is to drive the output $y(n)$ to a desired output response $y_{\text{d}}(n)$. During the k-th iteration the PA is driven by an input $u_k(n)$ which produces an output $y_k(n)$. The learning controller then uses the error observed between the desired and actual output $e_k(n) = y_{\text{d}}(n) - y_k(n)$ to modify the input signal that will be used during the next iteration $u_{k+1}(n)$. The learning algorithm is designed to ensure that the error $e_k(n)$ is reduced after each iteration. This process is repeated iteratively until the desired performance is reached.

The most important part in the design of an ILC scheme is the derivation of the learning algorithm. This is because that algorithm will control the

Table 3.1. Summary of ILC learning algorithms for PA linearization

Type	Algorithm	Learning matrix/gain
Gain-based	$\mathbf{u}_{k+1} = \mathbf{u}_k + \mathbf{G}(\mathbf{u}_k)^{-1}\mathbf{e}_k$	$\mathbf{G}(\mathbf{u}_k) = \mathbf{diag}\{\mathbf{y}_k/\mathbf{u}_k\}$
Linear	$\mathbf{u}_{k+1} = \mathbf{u}_k + \gamma\mathbf{e}_k$	$0 < \gamma < 2/J_{\max}$

convergence properties of the scheme. In Paper [A], we present the complete derivation of two learning algorithms for PA linearization purposes: the instantaneous gain-based ILC and the linear ILC algorithms. A summary of those algorithms is shown in Table 3.1.

In order to improve their convergence, the input signal used in the first iteration $u_1(n)$ must drive the output signal reasonably close to $y_{\mathrm{d}}(n)$, for PAs the initial input signal may be chosen as

$$u_1(n) = \frac{y_{\mathrm{d}}(n)}{g} \tag{3.5}$$

where g is a scaling factor that ensures that $u_1(n)$ does not exceed the PA maximum allowed input power level, for safe operation of the PA. A good choice of g may be the average gain g_{avg} of the amplifier at the desired average output power, which can be calculated from preliminary measurements.

The ILC scheme for the PA linearization can be summarized as follows:

Step 1) Select the desired PA output $\mathbf{y}_{\mathrm{d}}$

Step 2) Set $k = 1$ and let the input signal be $\mathbf{u}_1 = \mathbf{y}_{\mathrm{d}}/g_{\mathrm{avg}}$, where g_{avg} is the average gain of the amplifier at the desired average output power

Step 3) Apply the input $\mathbf{u}_k$ to the PA and measure the PA output $\mathbf{y}_k$

Step 4) Compute the error as $\mathbf{e}_k = \mathbf{y}_{\mathrm{d}} - \mathbf{y}_k$

Step 5) If the error satisfies the requirements, stop. Otherwise, go to the next step

Step 6) Compute the PA input signal for the next iteration $\mathbf{u}_{k+1}$ using any of the algorithms shown in Table 3.1

Step 7) Let $k = k + 1$ and go to Step 3

To show the linearization capabilities of this scheme, in Paper [A], the scheme was used to linearize a PA that is driven in high compression. The algorithm used was the linear ILC algorithm. The NMSE and ACPR values obtained with ILC were of -47.96 dB and -58.62 dBc, respectively. Fig 3.5 shows the evolution of the spectrum of the output signal after each iteration. Note that by using a simple algorithm as the linear ILC algorithm, ILC was able to eliminate all of the distortion introduced by the PA, as can be noticed from the spectrum plot, where the spectral regrowth reached the noise floor.

While ILC is a powerful technique to obtain the optimal signal that linearizes a PA, it is important to note that it uses the same desired output

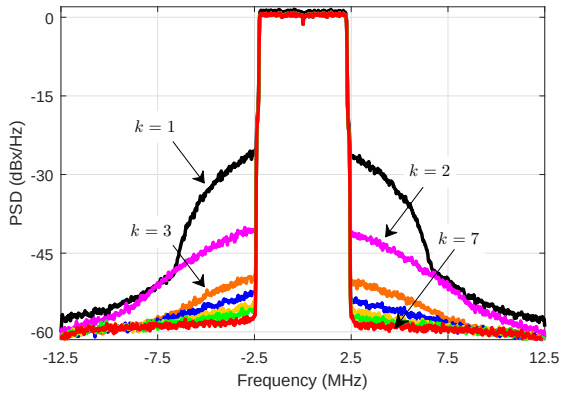


Figure 3.5. Power spectral density (PSD) of the measured PA output signal obtained using ILC at different iterations k

response y_{d} in every iteration of the system. For this reason, ILC cannot be directly used in practical linearization scenarios where the desired output from the PA is constantly changing. The true potential of ILC is that for the first time we can have access to the optimal signal that linearizes a PA/output signal from an optimal predistorter. This information can give us a better insight into the behavior of the pre-inverse of a PA and allows us to treat the problem of DPD as a behavioral modeling problem.

In the next two chapters, we will see how the ILC framework can be used in different aspects of the synthesis of digital predistorters. In Chapter 4, we will show how ILC can be used to derive model structures for digital predistorters; and in Chapter 5, we will see how we can use ILC in the parameter identification of predistorter models.

Chapter 4

Model Structures for Digital Predistortion

Selecting a model for the predistorter is the most important step in the synthesis of a digital predistorter. This is because the accuracy of that model will limit the linearity performance of the system. Over the years, different approaches have been used to select model structures for DPD applications. In early DPD works, predistorter models were derived by analytically inverting a forward model of the PA, which was a process that require complex derivations. In order to simplify the complexity of the DPD synthesis and thanks to the introduction of parameter identification techniques such as the indirect learning architectures, researchers opted to approximate PA inverse structures with models utilized to characterize their forward behavior, i.e., PA behavioral models [55]. The motivation for their choice was the idea that the inverse of a PA that presented memory should also be a nonlinear system with memory [55], in that way different PA behavioral models have been used for DPD purposes. The simplicity of the parameter extraction and the reasonable performance obtained by those models made this a popular approach as reflected in the literature [16, 43, 55–57].

Another approach that has prompted a lot of attention of the DPD community in recent years is the use of sparse approximation techniques to simplify the structure of DPD models. The main idea of those techniques is to take a general model that contains a large number of basis functions and use optimization algorithms to find an efficient subset that does not compromise the linearity performance. Sparse approximation is a mature field on its own and different existing algorithms have been applied to DPD, e.g., [58–61].

This chapter presents a new approach to derive model structures for digital predistorters using the concept of ILC. We begin this chapter explaining the idea behind this approach and then use it to derive a predistorter model structure based on a memory polynomial model.

4.1 Structured predistorter model derivation using iterative learning control

As demonstrated in Paper [A], ILC is a powerful technique capable of identifying the optimal predistorted signal that linearizes a PA. In Paper [B], we take advantage of the ILC concept to propose a new approach to derive model structures for predistorters. The idea is simple, assuming that a transfer function of the PA is known, ILC can be used to derive an approximate expression of the predistorted signal which can be used to identify basis functions for the predistorter model.

Consider a PA whose input and output relation is given by $\mathbf{y} = \mathbf{F}_{\text{PA}}(\mathbf{u})$, where $\mathbf{u} = [u(0), u(1), ..., u(N-1)]^T$ and $\mathbf{y} = [y(0), y(1), ..., y(N-1)]^T$. The linearization goal in ILC is to find an optimal predistorted signal that drives the output $\mathbf{y}$ to a desired linear output response which is defined here as $\mathbf{y}_{\text{d}} = G\mathbf{u}_1$, where G denotes the amplification gain and $\mathbf{u}_1$ the signal to be amplified.

During the first iteration, the PA is driven with an input $\mathbf{u}_1$, producing an error

$$\mathbf{e}_1 = \mathbf{y}_{\text{d}} - \mathbf{F}_{\text{PA}}(\mathbf{u}_1) = G\mathbf{u}_1 - \mathbf{F}_{\text{PA}}(\mathbf{u}_1) \tag{4.1}$$

Using the linear ILC algorithm shown in Table 3.1, the signal to be used in the second iteration is given by

$$\mathbf{u}_2 = \mathbf{u}_1 + \gamma\mathbf{e}_1. \tag{4.2}$$

Applying $\mathbf{u}_2$ to the PA and assuming that $\mathbf{F}_{\text{PA}}$ is continuous in the region of interest, the output in the second iteration will be given by

$$\begin{aligned}\mathbf{y}_2 &= \mathbf{F}_{\text{PA}}(\mathbf{u}_2) = \mathbf{F}_{\text{PA}}(\mathbf{u}_1 + \gamma\mathbf{e}_1). \\ &\approx \mathbf{F}_{\text{PA}}(\mathbf{u}_1) + \gamma J_{\mathbf{F}}(\mathbf{u}_1)\mathbf{e}_1\end{aligned} \tag{4.3}$$

where $J_{\mathbf{F}}(\mathbf{u}_1)$ is the Jacobian matrix of $\mathbf{F}_{\text{s}}$ with respect to $\mathbf{u}_1$.

Then the error in the second iteration will be given by

$$\begin{aligned}\mathbf{e}_2 &= \mathbf{y}_{\text{d}} - \mathbf{F}_{\text{PA}}(\mathbf{u}_2) \\ &= G\mathbf{u}_1 - \mathbf{F}_{\text{PA}}(\mathbf{u}_1) - \gamma G J_{\mathbf{F}}(\mathbf{u}_1)\mathbf{u}_1 + \gamma J_{\mathbf{F}}(\mathbf{u}_1)\mathbf{F}_{\text{PA}}(\mathbf{u}_1)\end{aligned} \tag{4.4}$$

Using again the linear ILC algorithm, the input signal to be used during the third iteration can be calculated as

$$\begin{aligned}\mathbf{u}_3 =& \mathbf{u}_2 + \gamma\mathbf{e}_2 \\ =& \big(1 + 2\gamma G\big)\mathbf{u}_1 - 2\gamma\mathbf{F}_{\text{PA}}(\mathbf{u}_1) - \gamma^2 G J_{\mathbf{F}}(\mathbf{u}_1)\mathbf{u}_1 + \gamma^2 J_{\mathbf{F}}(\mathbf{u}_1)\mathbf{F}_{\text{PA}}(\mathbf{u}_1)\end{aligned} \tag{4.5}$$

Using ILC, an expression of the predistorted signal as a function of the initial input signal (signal to be amplified) $\mathbf{u}_1$ can be derived. This expression

contains terms of the form

$$\mathbf{u}_1 \quad \mathbf{F}_{\mathrm{PA}}(\mathbf{u}_1) \quad J_{\mathbf{F}}(\mathbf{u}_1)\mathbf{u}_1 \quad J_{\mathbf{F}}(\mathbf{u}_1)\mathbf{F}_{\mathrm{PA}}(\mathbf{u}_1). \tag{4.6}$$

If the PA transfer function $\mathbf{F}_{\mathrm{PA}}$ is known, these terms can be used to find basis functions for the predistorter model.

4.2 Deriving a predistorter model based on the memory polynomial model

In Paper [B], the derivation method is used to derive a predistorter model assuming that the PA transfer function can be described using a memory polynomial (MP) model which is given by [17]

$$y(n) = \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} \sum_{m=0}^{M} a_{pm} u(n-m)|u(n-m)|^{p-1} \tag{4.7}$$

where P and M are the nonlinear order and memory depth of the model, respectively. This model choice is motivated by its simplicity and ability to characterize PAs exhibiting memory effects. However, this method could be applied to any other PA behavioral model.

After replacing (4.7) in (4.6) and rearranging the basis functions, the following model structure is derived

$$\begin{aligned} y(n) = & \sum_{\substack{p=1 \\ p \text{ odd}}}^{2P-1} \sum_{l=0}^{M} a_{pl} x(n-l)|x(n-l)|^{p-1} + \\ & \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} \sum_{m=M+1}^{2M} b_{pm} x(n-m)|x(n-m)|^{p-1} + \\ & \sum_{\substack{k=3 \\ k \text{ odd}}}^{P} \sum_{l=0}^{M} \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} \sum_{m=1}^{M} c_{pmkl} x(n-m-l)|x(n-m-l)|^{p-1}|x(n-l)|^{k-1} \end{aligned} \tag{4.8}$$

where P and M denotes the nonlinear order and memory depth of the MP forward model of the PA. This model is denoted as the memory polynomial based predistorter (MP-PD) model. Note that although this model structure share common basis functions with the MP model, it has a different structure.

Experimental results

The performance of the derived model was validated and compared to two models widely used in DPD, the MP model [17] and the GMP model [16] described in (4.7) and (2.4), respectively.

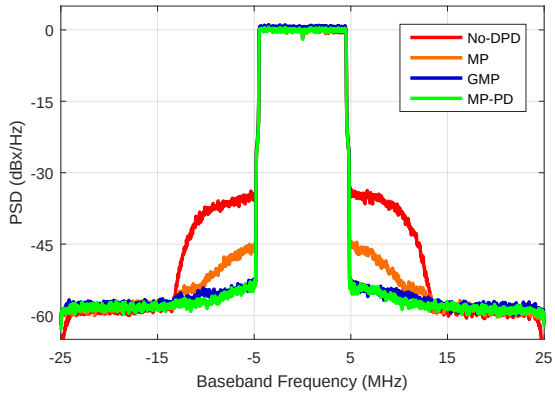


Figure 4.1. Power spectral density (PSD) of the output signal obtained before and after the PA was linearized with a memory polynomial (MP) model, a generalized memory polynomial (GMP) model and a memory polynomial based predistorter (MP-PD) model .

Table 4.1. Summary of linearization results

	Model Param.			**NMSE**	**ACPR**	**Coeff.**
	P	**M**	**L**	**(dB)**	**(dBc)**	**(#)**
w/o-DPD	-	-	-	-29.7	-37.6	-
MP	5	4	0	-41.7	-49.5	15
GMP	5	4	1	-44.9	-54.4	33
	5	4	2	-45.5	-55.6	49
MP-PD	3	4	-	-46.4	-55.7	55

The parameter settings of the MP-PD models were determined by estimating first a MP forward model of the PA. Different parameter settings were tested for the MP and GMP models. A summary of the best linearization results for all the models are summarized in Table 4.1 and the spectrum of the PA output signals obtained for all the models are shown in Fig. 4.1. Note that the MP-PD model achieved better NMSE performance than the other models tested, 4.7 dB better than the MP model and $\approx$ 1 dB better than the GMP model. In terms of ACPR, no substantial difference between the MP-PD and GMP is observed which can be explained by the limited dynamic range of the spectrum analyzer.

Chapter 5

Parameter Identification Techniques

Once a model has been chosen, the next step in the synthesis of a digital predistorter is to identify the parameters of that model. In PA behavioral modeling, because the input and output from the PA are known, the parameter estimation is a straightforward process. In DPD, however, because the optimal output from the predistorter is typically unknown, more advanced iterative techniques are required to identify the predistorter model parameters. This chapter discusses different techniques developed for this purpose. The chapter starts with a review of ILA. Thereafter, the issue of gain normalization in ILA is discussed and an alternative formulation of ILA is proposed. Next, DLA is reviewed. After that, identification techniques based on ILC are presented. The ILC-based DPD scheme and the adaptive ILC-based DPD scheme are discussed in Section 5.3 and 5.4, respectively. Finally, in Section 5.5, the idea of predistorter parameter identification using only one of the IQ components of the PA output signal is explained, and the real-valued ILC-based DPD scheme is introduced.

5.1 Indirect learning architecture

First introduced in [62] as a learning architecture to train neural network controllers and later adopted in [47] for DPD synthesis, ILA is currently the most widely known and used technique in DPD studies to identify the parameters of digital predistorters. A block diagram of ILA is depicted in Fig. 5.1. This technique is based on the inverse modeling approach, where a post-inverse of the PA is identified by using the PA output signal $y(n)$ to model the PA input $u(n)$. Once the post-inverse of the PA (also known as postdistorter) is identified, the parameters of the postdistorter are copied to an identical model which is used as the predistorter [13].

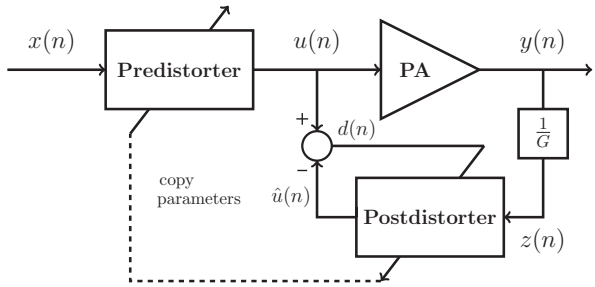


Figure 5.1. Indirect learning architecture principle. This technique identifies a postdistorter of the PA and uses it as predistorter.

Because ILA estimates a post-inverse rather than a pre-inverse, it focuses on minimizing the error between the PA input signal and the output of the predistorter, i.e., $d(n) = u(n) - \hat{u}(n)$ in Fig. 5.1. The minimization criteria used by ILA can be written in vector form as follows

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \left\| \mathbf{d} \right\| = \arg\min_{\boldsymbol{\theta}} \left\| F_{\mathrm{PD}}[\mathbf{x}, \boldsymbol{\theta}] - F_{\mathrm{PD}}[\mathbf{y}/G, \boldsymbol{\theta}] \right\| \tag{5.1}$$

where F_{PD} denotes the predistorter and postdistorter function. Under the assumption that the PA transfer function is one-to-one, ILA is based on the fact that if the error $d(n)$ vanishes, i.e., $d(n) = 0$ [38, 47] then

$$F_{\mathrm{PD}}[\mathbf{x}, \boldsymbol{\theta}] = F_{\mathrm{PD}}[\mathbf{y}/G, \boldsymbol{\theta}] \tag{5.2}$$

and consequently,

$$\mathbf{y} = G\mathbf{x}. \tag{5.3}$$

However, in general cases, $d(n)$ will not vanish completely. This may happen when the predistorter model structure used may not be sufficiently accurate or the parameters of the predistorter model may not be estimated perfectly. For instance, the authors of [63, 64] shown that if the parameters $\boldsymbol{\theta}$ are estimated using least squares (LS) due to the presence of measurement noise in the PA output signal $y(n)$, the parameters estimates converge to a biased solution. Despite those shortcomings, the ease of its identification has made it the most widely used parameter identification in the literature.

To overcome the noise-induced bias problem of ILA, a modified version of the ILA termed *model-based ILA* (MILA) was proposed in [63]. MILA reduces the effects of measurement noise by replacing the noisy measured output signal $y(n)$ with a noise-free model version $\hat{y}(n)$ in ILA.

5.1.1 Gain normalization issue in the indirect learning architecture

A critical issue encountered in ILA is that of the normalization gain selection [19,20]. The normalization gain G is the factor used to scale the output signal $y(n)$ to have the same power of the predistorter input signal $x(n)$, as shown in Fig. 5.1. In the literature, different ways to compute the normalization gain G have been proposed. Some of the most commonly known choices are:

1. The maximum gain of the PA G_{lin} [18]

2. The gain at the maximum targeted output power G_{peak} [19]

3. The gain adjustment technique based on power alignment [20], in which the normalization gain is adjusted in order to find a value that does not vary the average input power to the PA between iterations

The main issue with the normalization gain is that its selection affects the output power of the linearized amplifier. To better illustrate this, Fig. 5.2 presents the NMSE and ACPR results of the linearization of a Class AB PA using different normalization gain values. As noticed from this figure, the use of different normalization gain values produced output signals with different average power levels. Having to select a normalization gain adds an extra degree of freedom to the synthesis of the predistorter. Although the goal of DPD is to improve the linearity, it must also consider the output power of the PA. In order to properly evaluate the performance of a predistorter, it is required that the PA average output power obtained before DPD is not changed after DPD is applied [19]. To achieve this, the normalization gain must be carefully selected, which usually requires additional measurements and extra calibration efforts.

To overcome this issue, in Paper [C] a new formulation of ILA that eliminates the normalization gain block is proposed. The proposed ILA variant uses the desired PA output response $y_{\text{d}}(n)$ as input to the predistorter, as shown in Fig. 5.3. Since the output signal from the PA $y(n)$ and input signal to the predistorter $y_{\text{d}}(n)$ have similar power levels, no gain normalization at the PA output is required. Any power mismatch between those signals is directly handled by ILA. When ILA converges, the PA output is driven to $y(n) \approx y_{\text{d}}(n)$. Fig. 5.4 shows the linearization results obtained with the proposed ILA variant at different desired output power levels. Since the proposed ILA variant allows us to define the desired output signal before DPD is applied, the average output power obtained before and after DPD are the same for all the power levels tested. This simplifies the identification process since the average output power and linearity of the PA do not longer depend on the selection of a normalization gain.

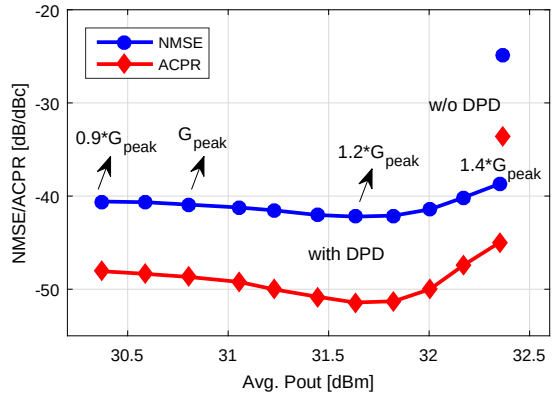


Figure 5.2. NMSE and ACPR results versus average output power obtained using different normalization gain values. Note that the use of different normalization gain values produce output signal with different average powers.

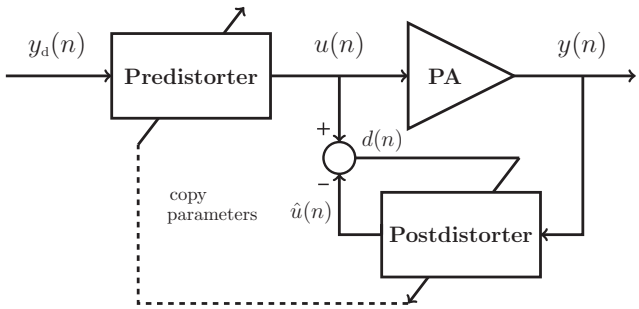


Figure 5.3. Proposed variant to the indirect learning architecture. Unlike the conventional indirect learning architecture, this variant eliminates the normalization gain and uses the desired PA output response as input to the predistorter.

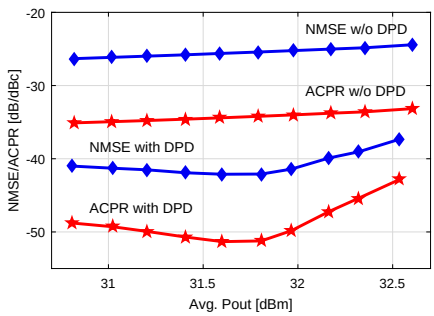


Figure 5.4. NMSE and ACPR results versus average output power obtained using the proposed ILA variant. Note that the average output power obtained before DPD does not change after DPD.

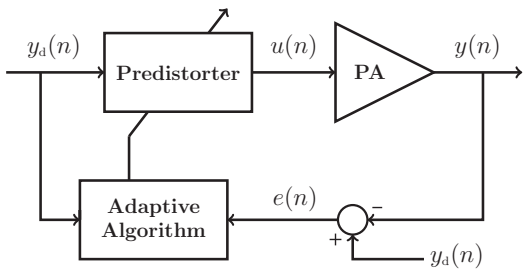


Figure 5.5. Direct learning architecture principle. This technique uses complex algorithms to directly estimate the parameters that minimize $e(n)$.

5.2 Direct learning architecture

The direct learning architecture (DLA), illustrated in Fig. 5.5, is a parameter identification technique that has grown in popularity in the last years. This is because unlike ILA, DLA focuses on minimizing the error between the desired and the actual PA output signals, i.e., $e(n) = y_{\mathrm{d}}(n) - y(n)$. The minimization criteria used in DLA can be written as follows

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \big\|\mathbf{e}\big\| = \arg\min_{\boldsymbol{\theta}} \big\|\mathbf{y}_{\mathrm{d}} - F_{\mathrm{PA}}\big[F_{\mathrm{PD}}[\mathbf{y}_{\mathrm{d}}, \boldsymbol{\theta}]\big]\big\|, \tag{5.4}$$

where F_{PA} and F_{PD} denote the transfer function of the PA and predistorter, respectively.

Different ways to implement DLA have been proposed, the most commonly known are: the model-based DLA, where the algorithm is run over a sufficiently accurate PA forward model [65–69], and the closed loop estimator, where the algorithm is run directly on the PA [70, 71].

The model-based DLA is implemented in two steps [65, 66]. First a forward model of the PA is identified. Once the model is obtained, complex optimization algorithms are used to estimate, through iterative processing, the predistorter parameters that minimize the criteria shown in (5.4). Once the nonlinear algorithm finds a solution, the parameters are used to generate a predistorted signal $u(n)$ that is applied to the real PA. This process is repeated iteratively until the predistorter-PA system converges to the best possible solution. Various model-based DLA algorithms have been proposed in the literature [65–67], unfortunately they are generally complex in structure, computationally expensive and present slow convergence.

The closed loop estimator is a DLA technique that has gained popularity in recent years because it does not rely on a PA forward model as previous DLA implementations did [70, 71]. The operation of this technique is similar to the model-based DLA described before with the difference that the algorithm is run directly on the PA. The update algorithm used in the closed loop estimator

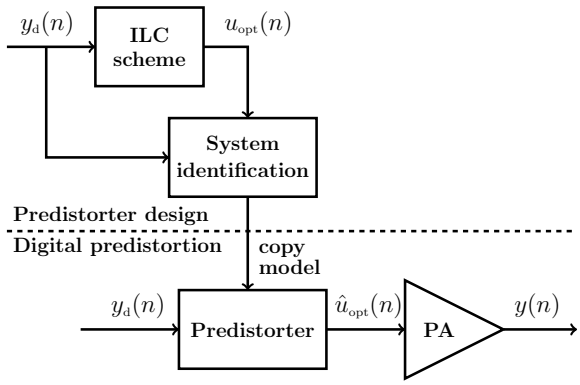


Figure 5.6. Block diagram of the iterative learning control based digital predistortion technique. In this technique, the ILC scheme described in Chapter 3 is first used on the power amplifier to identify the optimal predistorted signal $u_{\text{opt}}(n)$ that drives the PA to the desired linear output response $y_{\text{d}}(n)$. After that, $y_{\text{d}}(n)$ and $u_{\text{opt}}(n)$ are used to identify a predistorter model.

looks as follows

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \beta(\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{e}_k \tag{5.5}$$

where the subindex k indicates the iteration number and $\beta < 1$ is a step size.

5.3 Iterative learning control based digital predistortion

Introduced in Paper [A], the ILC-based DPD (ILC-DPD) scheme is an identification technique that combines the ILC scheme presented in Chapter 3 with standard estimation techniques to identify the parameters of a predistorter model, as depicted in Fig. 5.6. The idea is simple, since ILC can identify the optimal predistorted signal $u_{\text{opt}}(n)$ that linearizes a PA but does not provide a predistorter model, ILC-DPD uses that optimal predistorted signal and the desired output response to estimate a model for the predistorter, where $y_{\text{d}}(n)$ is used as model input and $u_{\text{opt}}(n)$ as model output.

Note that because ILC requires the same desired output signal to estimate the optimal PA input excitation, ILC-DPD is suited for offline linearization and research purposes. For instance, it can be used to select suitable model structures and settings for predistorters, because once the optimal predistorted signal is estimated, several structures and settings can be easily tested. This is opposed to ILA and DLA which require several iterations on the PA to evaluate the performance of each model.

Experimental results

In Paper [A], the linearization capabilities of ILC and ILC-DPD were evaluated and compared against ILA and the model-based DLA in two experimental scenarios. In the first scenario, their performance was evaluated when the output signal presented different levels of signal-to-noise ratio (SNR), while in the second scenario their performance was evaluated when the PA was in high compression. The algorithm used in the model-based DLA was the nonlinear least squares algorithm solver from MATLAB. The ILC algorithms used in the first and second scenario were the linear ILC algorithm and the gain-based ILC algorithm, respectively.

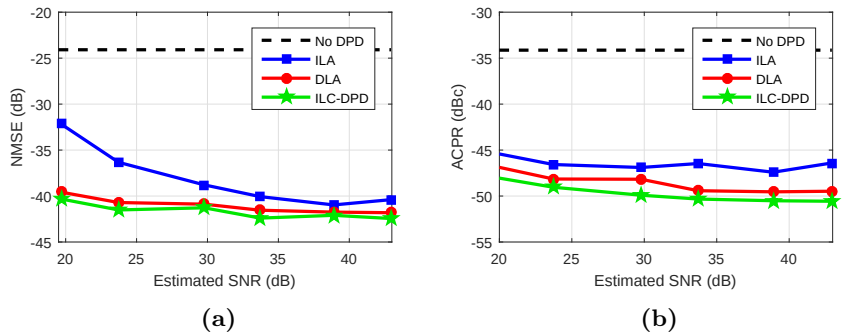
Fig. 5.7 shows the NMSE and ACPR results obtained in the first scenario. Compared to ILA, ILC-DPD achieved lower NMSE and ACPR values for all the SNR values tested. Note also that the performance improvement obtained using ILC-DPD is more significant at low SNRs where the ILA performance degrades at a much faster rate. This degradation is associated to the effects that the measurement noise has on the ILA estimates. Compared to the model-based DLA, ILC-DPD achieved similar linearity performance. However, for the model-based DLA, that performance came at the cost of a more complex identification process as it was described in Section 5.2. The algorithm used in the model-based DLA required hundreds of iteration to estimate the parameter of the predistorter, while the algorithm used in ILC-DPD required only 7 iterations to find the optimal predistorted signal and used standard modeling techniques to estimate the predistorter parameters.

Table 5.1 and Fig. 5.8 show the NMSE, ACPR, EVM and output spectrum results obtained in the second scenario. Note of all the techniques, ILC achieved the lowest NMSE and ACPR values which indicates that ILC can successfully identify the optimal predistorted signal that linearizes the PA. This can also be observed in Fig. 5.8 where the output spectrum obtained with ILC reached the noise floor. Note also that ILC-DPD achieved lower NMSE and ACPR values than ILA and similar values to DLA.

These results shown ILC-DPD is less sensitive to measurement noise than ILA and can provide better linearity performance when the PA nonlinearities are strong. In addition, ILC-DPD can achieve similar performance than DLA but with a simpler identification process.

5.4 Adaptive iterative learning control based digital predistorter

Since ILC requires the same desired output response to obtain the optimal predistorted signal that linearizes a PA, ILC-DPD is not suitable for real-time DPD applications. To overcome this problem, in Paper [D], the adaptive ILC-based DPD scheme is presented. The operation principle of this technique is similar to ILC-DPD with the difference that it updates the parameter of



(a) **(b)**

Figure 5.7. (a) NMSE and (b) ACPR versus estimated signal-to-noise ration (SNR) obtained without DPD, ILA, the model-based DLA and ILC-DPD.

Table 5.1. Summary of the linearization results obtained when the PA was driven under high compression

DPD type	Ident. NMSE (dB)	Ident. ACPR (dBc)	Valid. NMSE (dB)	Valid. ACPR (dBc)	Valid. EVM (dB)
w/o-DPD	-17.97	-32.42	-18.05	-32.27	-18.51
ILC	-47.96	-58.62	-	-	-38.43
ILC-DPD	-41.52	-50.16	-41.02	-50.17	-37.21
ILA	-39.85	-48.68	-39.33	-49.01	-36.03
DLA	-41.81	-50.94	-41.35	-50.94	-37.21

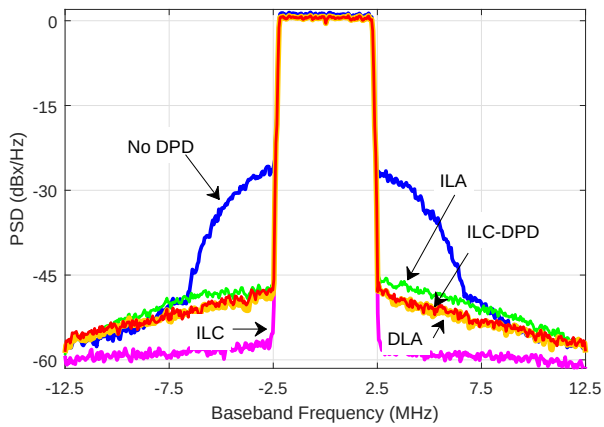


Figure 5.8. Power spectral density (PSD) of the measured PA output signal obtained without DPD, and after applying ILC, ILA, model-based DLA, and ILC-DPD for a Class B PA driven under high compression.

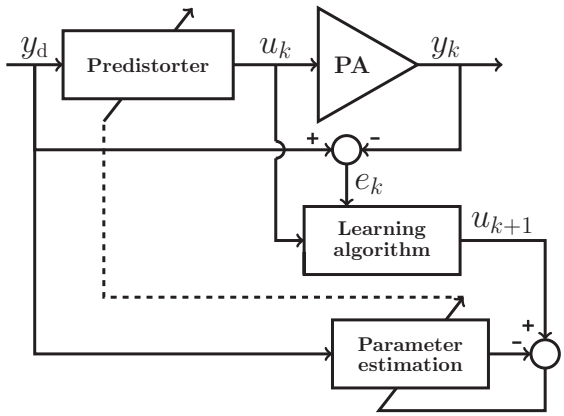


Figure 5.9. Block diagram of the adaptive iterative learning control based DPD scheme

the predistorter after each iteration, thereby removing the requirements of a repeated desired output signal.

The block diagram of the adaptive ILC-DPD is depicted in Fig. 5.9. During the k-th iteration, the desired output signal $y_d(n)$ is fed to the predistorter generating the predistorted signal $u_k(n)$ which is sent to the PA. Then, any of the learning algorithms presented in Table 3.1 is used to create a new input signal $u_{k+1}(n)$. Because the learning algorithms were derived to ensure that if $u_{k+1}(n)$ was sent directly to the PA, the PA output will be closer to the desired response $y_d(n)$; in the adaptive ILC-DPD, $u_{k+1}(n)$ is used in combination to $y_d(n)$ to estimate the parameters of the predistorter model, where $y_d(n)$ is used as model input and $u_{k+1}(n)$ as model output. This process is repeated iteratively until a given performance is reached. Experimental results with this technique are presented in the following section.

5.5 Parameter identification using real-valued output data

The need to increase the bandwidth of the communication signals to support the increasing demands for higher data rates creates new challenges in the implementation of DPD systems. One of the most crucial ones is the high sampling rates required in the DPD feedback receiver. In order to effectively linearize a PA, a DPD feedback receiver should cover a span that is typically five times the communication signal bandwidth [27]. This means that as the bandwidth of the communication signal continue to increase, so does the sampling speed of the analog to digital converters (ADC). This represents a problem because ADCs are expensive and power hungry components. Over the years, different approaches have been proposed to reduce the requirements

of ADCs in DPD feedback receivers. Among the most commonly known are the use of undersampling [23–26] and band-limited modeling [27, 28].

In Paper [D], a novel approach to relax the requirements of ADCs is presented. Instead of acquiring both IQ components of the PA output signal, the proposed approach only requires the acquisition of one of the IQ components. This approach combines an estimation approach that allows us to identify the parameters of a model using either the I or Q component of the model output with an adaptive DPD parameter identification technique.

5.5.1 Least squares estimation using real-valued output data

Consider two sets of complex-valued data, $\mathbf{u} = [u(0), u(1), \ldots, u(N-1)]^T$ and $\mathbf{y} = [y(0), y(1), \ldots, y(N-1)]^T$ whose relationship can be characterized using a linear model of the form

$$\mathbf{y} = \mathbf{H}\boldsymbol{\theta}, \tag{5.6}$$

where $\mathbf{H} = \mathbf{H}(\mathbf{u})$ is the regression matrix containing the basis functions of the model and $\boldsymbol{\theta}$ is a vector containing the model parameters. Using least squares, those parameters are generally calculated by [45]

$$\hat{\boldsymbol{\theta}} = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{y}, \tag{5.7}$$

that is, using both IQ components of $\mathbf{u}$ and $\mathbf{y}$. In Paper [D], it is shown that to estimate those parameters it is sufficient to have knowledge of both IQ components of the model input $\mathbf{u}$ and only one of the IQ components of the model output $\mathbf{y}$.

Defining $\boldsymbol{\Phi}$ as a vector containing the real and imaginary parts of the model parameters $\boldsymbol{\theta}$

$$\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\theta}_r \\ \boldsymbol{\theta}_i \end{bmatrix} \tag{5.8}$$

where the subscript r denotes the real part $\mathrm{Re}[\cdot]$, and the subscript i denotes the imaginary part $\mathrm{Im}[\cdot]$. It is shown that $\boldsymbol{\Phi}$ can be estimated using either of the following expressions

$$\hat{\boldsymbol{\Phi}} = (\mathbf{M}_r^T\mathbf{M}_r)^{-1}\mathbf{M}_r^T\mathbf{y}_r, \tag{5.9}$$

$$\hat{\boldsymbol{\Phi}} = (\mathbf{M}_i^T\mathbf{M}_i)^{-1}\mathbf{M}_i^T\mathbf{y}_i. \tag{5.10}$$

with $\mathbf{M}_r = \begin{bmatrix} \mathbf{H}_r & -\mathbf{H}_i \end{bmatrix}$ and $\mathbf{M}_i = \begin{bmatrix} \mathbf{H}_i & \mathbf{H}_r \end{bmatrix}$. If (5.9) is used, only the real part or I component of the model output signal $\mathbf{y}$ is required. Alternatively, if (5.10) is chosen, only the imaginary part or Q component of $\mathbf{y}$ is actually needed.

Because this approach uses half the information of the model output signal $\mathbf{y}$, the estimator is more sensitive to noise. This does not represent a major issue if the output signal presents a large SNR, as is the case in most DPD systems. In case the SNR is low enough to affect the performance, this could be recovered by increasing the number of samples.

5.5.2 Real-valued iterative learning control based digital predistortion

To take advantage of the estimation approach previously described to reduce the requirements of ADCs, a parameter identification technique in which the signal used as model output is a function of the PA output signal is required. In Paper [D], that approach is used in combination with the adaptive ILC-DPD scheme described in Section 5.4. The technique is termed the real-valued ILC-DPD (RILC-DPD) . Note however that it could be used in combination with any identification technique that fulfills that property, for instance, it could be used with DLA or the model-based ILA [63].

Since only one of the IQ components of the model output $u_{k+1}(n)$ is required, a modified version of the linear ILC algorithm that considers only the real part of $u_{k+1}(n)$ is used

$$\mathbf{u}_{k+1,r} = \mathbf{u}_{k,r} + \gamma \mathbf{e}_{k,r} \tag{5.11}$$

where the subscript r denotes the real part and γ is the learning gain which can be also be chosen using only the I component of the PA output signal as is described in Paper [D].

Experimental results

The proposed RILC-DPD scheme was evaluated and compared against the conventional adaptive ILC-DPD which uses complex-valued measurement output data, and ILA [72] in two experimental scenarios. In the first scenario, the performance of the three techniques were assessed using a sufficiently large linearization bandwidth, i.e., five times the bandwidth of the signal. In the second scenario, the three techniques were used in combination with the bandlimited modeling approach proposed in [27] to evaluate their performance in a limited linearization bandwidth scenario.

The output spectrum obtained with all the identification techniques in the first linearization scenario is shown in Fig. 5.10, and a summary of the linearization results is presented in Table 5.2. Despite using only one of the IQ components, RILC-DPD provided similar linearization performance as the adaptive ILC-DPD scheme. This can also be appreciated from Fig. 5.10, where the spectrum of both techniques overlap each other. The results also show that RILC-DPD and the adaptive ILC-DPD can provide better linearity performance than ILA.

Table 5.2. Linearization results obtained in the sufficiently large linearization bandwidth scenario

	GMP			NMSE	ACPR	EVM
	P	**M**	**L**	**(dB)**	**(dBc)**	**(dB)**
w/o-DPD	-	-	-	-21.90	-29.67	-24.06
ILC-DPD	9	3	1	-40.60	-49.48	-42.90
RILC-DPD	9	3	1	-40.69	-49.36	-43.20
ILA	9	3	1	-37.47	-44.80	-40.15

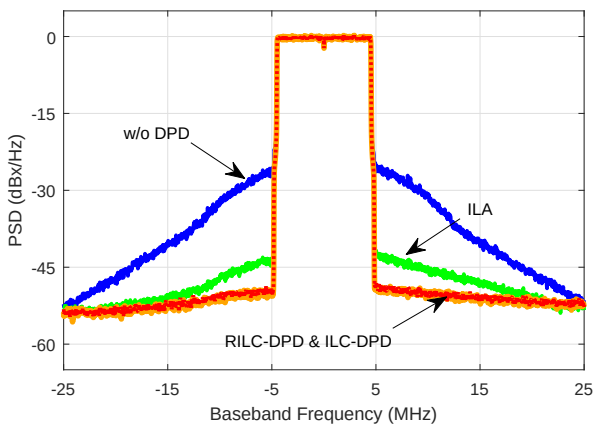


Figure 5.10. Spectrum of the PA output signal obtained without DPD and after using the proposed RILC-DPD, the conventional complex-valued ILC-DPD, and ILA. Note that the spectrum of the output signals obtained with RILC-DPD and ILC-DPD overlap each other.

Table 5.3. Linearization results obtained in the limited linearization bandwidth scenario

	Model (P,M,L)	**BW MHz**	**NMSE (dB)**	**ACPR (dBc)**	**EVM (dB)**
w/o DPD	-	100	-22.84	-30.59	-24.67
	-	60	-22.82	-30.61	-24.60
RILC-DPD	GMP(9,3,1)	100	-40.36	-50.35	-42.56
	GMP(9,3,1)	60	-38.57	-45.74	-40.75
	BL-GMP(9,3,1)	60	-41.10	-50.47	-42.41
ILC-DPD	GMP(9,3,1)	100	-40.97	-50.23	-43.85
	GMP(9,3,1)	60	-39.20	-45.78	-41.78
	BL-GMP(9,3,1)	60	-41.73	-50.29	-43.29
ILA	GMP(5,3,2)	100	-35.03	-41.75	-39.17
	GMP(5,3,2)	60	-34.66	-40.40	-38.21
	BL-GMP(5,3,2)	60	-35.53	-41.03	-38.95

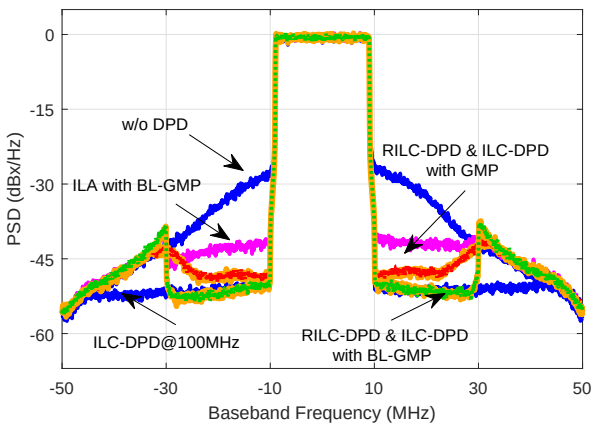


Figure 5.11. Spectrum of the measured output signal obtained without DPD and after using the proposed RILC-DPD, ILC-DPD, and ILA with a linearization bandwidth of 60 MHz and GMP and band-limited (BL)-GMP models. For reference the spectrum of the output signal after using ILC-DPD with a linearization bandwidth of 100 MHz. Note that RILC-DPD with a BL-GMP model provide similar spectral leakage reduction as ILC-DPD with a linearization bandwidth of 100MHz.

In the second scenario, the PA was linearized using RILC-DPD, the adaptive ILC-DPD and ILA in three experimental cases:

1. Using a GMP model with a linearization bandwidth of 100 MHz
2. Using a GMP model with a linearization bandwidth of 60 MHz
3. Using a BL-GMP model [27] with a linearization bandwidth of 60 MHz

Fig. 5.11 shows the spectrum of the output signals obtained with the different identification techniques, in the second and third cases. Table 5.3 summarizes the linearization performance obtained with the different identification techniques. For RILC-DPD and ILC-DPD, the settings of the GMP and BL-GMP models were: nonlinear order $P = 9$, memory depth $M = 3$ and cross-term length $L = 1$. The same settings were tested for ILA, but excessive amplitude expansion of the PA input signal made it unfeasible to test its performance without damaging the PA. Several other settings were tested for ILA, the best performance was achieved with $P = 5$, $M = 3$, and $L = 2$.

Comparing the performance in the three cases, note that good performance is achieved with a linearization bandwidth of 100 MHz. The performance is degraded when the linearization bandwidth was reduced to 60 MHz and a GMP model was used, but the performance can be recovered using a BL-GMP model instead. The degradation going from 100 MHz to the 60 MHz case is caused by aliasing distortion introduced in the basis functions of the model as a result of using a linearization bandwidth of only three time the channel bandwidth.

Comparing the identification techniques, note that in all the cases, RILC-DPD achieved similar performance to the adaptive ILC-DPD that uses both IQ components of the output signal. This can be observed in Fig. 5.11, where the output signal spectrum of both techniques overlap each other. Compared to ILA, RILC-DPD obtained better linearity performance in all the cases. The results demonstrate that RILC-DPD can be sucessfully combined with the band-limited modeling approach to not only reduce the number of ADCs but also their speed requirements.

Chapter 6

Limits on the linearity performance in radio transmitters

In DPD, the validation of a proposed model or identification technique is generally done by comparing them against existing approaches in the literature, hoping to achieve better performance. While this seems reasonable, it does not take into consideration how far such performance is from the ideal one, to see if further improvement is possible. This chapter discusses the performance limits in the linearization of PAs, particularly on the limits of the NMSE. The idea is to provide a better understanding of the behavior of the NMSE of a linearized PA with respect to the output power, which is something that is often overlooked in DPD studies.

6.1 Lower bound for the normalized mean square error

As described in Section 3.3, the NMSE is one of the most common performance metrics used in DPD studies to evaluate the linearity of PAs. In Paper [E], a lower bound for the NMSE performance of a linearized PA is derived.

The system consider in the analysis is depicted in Fig. 6.1. The desired output signal y_{d} is passed through the predistorter-PA system generating the output signal y. When the signal is acquired, measurement noise w is added to the output from the PA resulting in the measured output signal $y_w = y + w$. In the analysis, y_{d} is considered to be complex Gaussian distributed with zero mean and variance $2\sigma^2$, i.e., $y_{\mathrm{d}} \sim \mathcal{CN}(0, 2\sigma^2)$ and the measurement noise is a zero-mean complex additive white Gaussian noise (AWGN) with variance σ_w^2. In a perfect linearization scenario, the predistorter should be able to perfectly

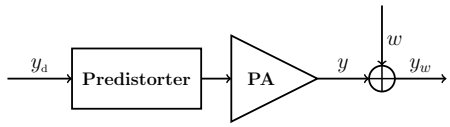


Figure 6.1

compensate all the distortion introduced by the PA up to the saturation point. This means that the input-output relation of predistorter-PA system could be described by [73]

$$y = \begin{cases} y_{\mathrm{d}} & ,|y_{\mathrm{d}}| \leq a \\ a\exp(j\angle y_{\mathrm{d}}) & ,|y_{\mathrm{d}}| > a \end{cases} \tag{6.1}$$

where $|.|$ denotes the absolute value, a denotes the PA output saturation point, and $\angle y_{\mathrm{d}}$ denotes the phase of y_{d}.

Assuming the conditions mentioned above, it is shown in Paper [E] that the lower bound for the NMSE is given by

$$\mathrm{NMSE} = \frac{2\sigma^2\exp\left(\frac{-a^2}{2\sigma^2}\right) - 2a\sigma\sqrt{2\pi}Q\left(\frac{a}{\sigma}\right) + \sigma_w^2}{2\sigma^2}. \tag{6.2}$$

where $Q(.)$ denotes the Q-function In a similar way, the average output power from the PA delivered to a load impedance $R = 50\Omega$ can be calculated by

$$P_{\mathrm{out,avg}} = \frac{1}{R}(\sigma^2 - \sigma^2\exp\left(\frac{-a^2}{2\sigma^2}\right)) + \frac{\sigma_w^2}{2R}. \tag{6.3}$$

From these expression, it can be noticed that to analyze the behavior of the NMSE as a function of the average output power from the PA only three parameters are required: the variance of the desired output signal, the noise variance and the saturation point of the PA.

6.2 Simulation and experimental results

The analytical framework was validated through numerical simulations. In addition, measurements on a PA were performed in order to verify if the derived NMSE lower bound could be attained with existing linearization techniques. The techniques considered were ILC, the adaptive ILC-DPD and ILA.

Two separate sets of OFDM signals were used, each set consisted of 10^6 samples at a sampling rate of 200 MHz. The signals had a PAPR of ≈ 11 dB. The predistorter model used in ILC-DPD and ILA was the vector-switched generalized memory polynomial (VS-GMP) model [43]. The phase noise in the measured output signal was estimated and compensated for using a Kalman filter which tracks the slow variations in the phase transfer function of the PA [74].

Fig. 6.2 presents the theoretical, simulated and experimental NMSE results before and after phase noise compensation. Fig. 6.3 shows the simulated and experimental ACPR results. Note that the theoretical NMSE results match the simulated ones, which proves the accuracy of the NMSE lower bound derivation.

From the theoretical results, it can be noticed that the NMSE presents two distinct behaviors. At low output power levels, the NMSE decreases at a rate of 1 dB per 1 dB increase in the average output power. In that region, the NMSE is limited mainly by the variance of the measurement noise. As the output power level increases above 28 dBm, the peaks of the PA output signal starts to be clipped and the NMSE starts to degrade. At the beginning, the degradation is small because only few samples are clipped but eventually the NMSE is dominated by the distortion introduced by the saturation of the output signal. Note also that although the NMSE starts to degrade at about 28 dBm, for the PA tested acceptable NMSE and ACPR performance, i.e., NMSE lower than -35 dB and ACPR lower than -45 dBc, can still be obtained up to a power level of 30.5 dBm, which means that up to 2.5 dB more output power could be gained if the PA is allowed to work in that region.

The results also show that at low output power levels, all the linearization techniques could reach the NMSE lower bound after phase noise compensation. But the results deviated from the ideal ones as the output power level was increased. The lowest NMSE performance was obtained by ILC. Because ILC is nonparametric, it was not limited by the accuracy of a predistorter model and was able to compensate for most of the nonlinear distortion at the PA output. However, ILC was not able to reach all the way to the NMSE lower bound. The residual distortion may be caused due to limitations of the linear ILC algorithm which uses a scalar learning gain, and for some effects that produced a slightly different behavior in the PA every time this was operated. ILC can only compensate for effects that are repeatable.

With respect to ILC-DPD and ILA, both achieved the same NMSE up to an average to an average output power level of $\approx$ 26.4 dBm, but for higher output power levels, the results obtained with ILA degraded at a faster rate than ILC. ILC-DPD achieved better performance than ILA, but it is far from reaching the NMSE lower bound. The performance degradation of ILC-DPD and ILA compared to ILC was caused by the limited accuracy of the predistorter model. But the higher degradation of ILA was caused by the nature of its identification process. ILA estimates a post-distorter and assumes that it can be used as predistorter. This assumption works fine when the PA nonlinearity is not too severe, as we can see from the results for power levels below 26.4 dBm, but as the PA is driven more into compression, that assumption is not longer valid.

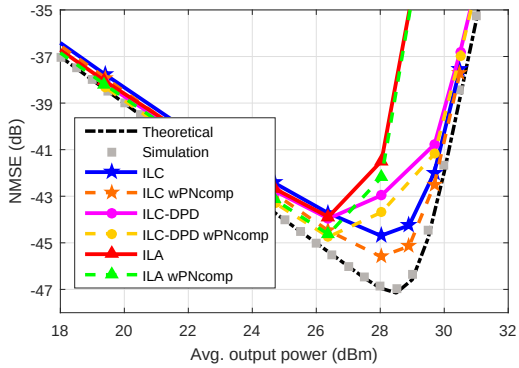


Figure 6.2. Experimental NMSE results obtained after linearization and phase noise compensation

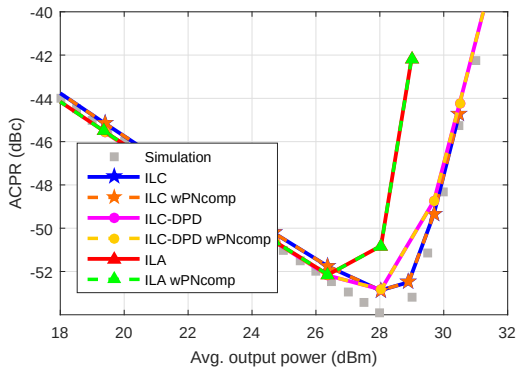


Figure 6.3. Experimental ACPR results obtained after linearization and phase noise compensation

Chapter 7

Conclusions and summary of appended papers

7.1 Conclusions

High efficiency and linear PAs are essential components in wireless communication systems. There is however a tradeoff between efficiency and linearity. In order to fulfill both requirements, DPD is often used. This thesis has contributed to different aspects of the linearization of PAs through DPD.

An issue encountered in DPD has been that the output of the signal from an optimal predistorter is unknown. To overcome this problem, in this work, the concept of ILC for the linearization of PAs was introduced. Instead of focusing on identifying the parameters of a predistorter model, ILC focuses first on estimating the optimal predistorted signal that drives a PA to desired linear output response. Experimental results showed that, even for the most difficult cases, ILC can successfully find such optimal predistorted signal. But the significance of ILC is more than a way to identify the optimal signal that linearizes a PA. To the author's opinion, this technique provides a mathematical foundation to help analyze the problem of DPD. This was demonstrated in Chapter 4, where ILC was used to develop a novel technique to derive models for digital predistorters. It was shown that ILC can be used to find an analytically expression of a predistorted signal which can help us identify proper basis function for predistorters. The technique was used to derive a model structure based on the MP model and its performance was compared to two commonly used models. The performance showed that the model derived with the proposed technique could achieve lower NMSE performance than other models. While physical interpretations of PA behavioral models have been proven for PA forward behavior, this is not the case for their inverse behavior. The proposed derivation technique could be used in combination with such PA forward models to help identify key basis functions for DPD models.

The concept of ILC can also be used in the identification of the parameters of the predistorter as it was shown in Chapter 5. Although ILC does not provide a predistorter model per se, the optimal predistorted signal is still useful in DPD studies. This signal can give us an insight to the characteristics of the PA inverse behavior and can be used to find proper model structures and model settings for the predistorter. This approach may be particularly useful in linearization scenarios where other identification techniques fail, because it allows us to attack the problem of model structure design and parameter estimation separately. The ILC framework can also be used in adaptive linearization scenarios. The ILC algorithms derived in this work can be easily integrated in online linearization scenarios by updating the parameter of the predistorter after each iteration.

In DPD, the parameter identification is done using complex-valued baseband equivalent signals of the PA RF input and output signals. In Chapter 5, we introduced the idea of performing the parameter identification using only one of the IQ components of the PA output. By combining an estimation approach that allows to estimate the parameter of a model using either the I or Q component of the model output with the adaptive ILC based scheme, it was shown that similar linearity performance to identification techniques that use both IQ components can be obtained. Since only one of the IQ components needs to be acquired, this technique may help to reduce the number of ADCs used in DPD feedback receivers, whose speed requirements, power consumption escalate with the increase of the communication signal bandwidths.

The problem of gain normalization in DPD schemes was also discussed in Chapter 5. It was shown that the use of a normalization gain block in DPD schemes increases the complexity of the parameter identification process. PAs have different gain values at different input power levels which makes the idea of gain selection confusing. In addition, the normalization gain selection does not only affect the linearity performance but also the output power of the linearized PA, therefore it adds an extra degree of freedom which must be carefully selected. We have shown that by reformulating the predistorter goal to not only include the linearity but also the desired output power, the identification process may be simplified.

In Chapter 6, the performance limits in the linearization of PAs are discussed. In DPD studies, the performance of a technique is evaluated by comparing their performance against existing techniques. The goal is often to obtain lower NMSE, EVM or ACPR values without really knowing how low is good enough or whether the minimum achievable NMSE has been already reached. To help fill this gap, we derived a closed-form expression for the minimum NMSE that could be obtained in a linearization scenario. Three linearization techniques were tested: ILC, ILA and ILC-DPD. The experimental results shown that at low output power levels, all the techniques could reach the NMSE lower bound, but at higher output power levels, their performance degraded from the bound. The performance degradation for ILA and ILC-

DPD was mainly due to the limited accuracy of the predistorter model. For ILC, the degradation may be caused due to limitations of the ILC algorithm used and for some effects that produced a slightly different behavior in the PA every time this was operated. ILC can only compensate for effects that are repeatable.

7.2 Future work

There are different sidelines of research arising from the work presented in this thesis which can be pursued. Because of the simplicity of the algorithms used in ILC, ILC provides a way to mathematically analyze the problem of DPD as was never done before. It would be interesting to mathematically analyze the linearization of difficult PAs, e.g., dual-input Doherty or outphasing PAs, to get an idea of how the predistorted signal should look like and to derive proper linearization schemes.

In this work we have shown the ILC-based DPD schemes can provide better linearity performance than ILA in SISO linearization cases. Therefore, it would also be interesting to extend the ILC scheme and the ILC-based DPD schemes proposed here for MIMO linearization to see whether it can also provide improved linearity performance.

7.3 Summary of appended papers

A summary of the appended papers are shortly described in the following.

Paper A: Iterative Learning Control for the Linearization of Power Amplifiers

In this paper, we propose a new technique to identify the parameters of a digital predistorter based on iterative learning control (ILC). Instead of focusing on identifying the predistorter parameters, the technique proposed here first uses an iterative learning algorithm to identify the optimal power amplifier (PA) input signal that drives the PA to the desired linear output response. Once the optimal PA input signal is identified, the parameters of the predistorter are estimated using standard modeling approaches, e.g., least squares. To this end, in this paper we present a complete derivation of an ILC scheme suitable for the linearization of PAs which includes convergence conditions and the derivation of two learning algorithms. Based on experimental results, it is shown that even for the most difficult cases, the proposed ILC scheme can successfully linearize the PA. The results also shown that the proposed parameter identification technique is less sensitive to measurement noise than ILA and can provide better linearity performance than DLA.

The author of this thesis was the main contributor of this paper, performed the experimental work and manuscript writing. The measurement results were

presented according to the insight given by the co-authors. The manuscript was reviewed and refined by the co-authors.

Paper B: Structured Digital Predistorter Model Derivation using Iterative Learning Control

In this paper, we present a novel approach to derive model structures for digital predistorters based on ILC. The ILC concept is used to derive an analytical expression of the predistorted signal which is used to identify basis functions for predistorter models. The proposed derivation approach is used to derive a predistorter model based on the memory polynomial model. The experimental results showed that the derive model can obtain better linearity performance than conventional models used in DPD.

The author of this thesis was the main contributor of this paper, performed the experimental work and manuscript writing. The manuscript was reviewed and refined by the co-authors.

Paper C: A New Variant of the Indirect Learning Architecture for the Linearization of Power Amplifiers

This paper investigates the effects that the normalization gain has on the linearization of PAs through ILA. In addition, it proposes a reformulation of the ILA that eliminates the need of the normalization gain. Experimental results show that the selection of the normalization gain affects the average output power and linearity performance of the linearized PA. If the normalization gain is not chosen correctly, the average output power of the linearized PA will differ from the average output power obtained before DPD. It is experimentally shown that the proposed ILA variant can alleviate that problem maintain the same average output power before and after DPD. Consequently the proposed ILA variant simplifies the linearization process and allows proper evaluation of the DPD performance

The author of this thesis was the main contributor of this paper, performed the experimental work and manuscript writing. The manuscript was reviewed and refined by the co-authors.

Paper D: Digital predistortion parameter identification technique using real-valued measurement output data

In this brief, we present a novel parameter identification technique that requires the acquisition of one of the IQ components of the PA output signal. To this end, we derive a technique that allows us to estimate the parameters of a model using only one of the IQ components of the model output. Based on experimental results, it is shown that the proposed parameter identification technique can provide similar linearization capabilities as its complex-valued

counterparts. Since the proposed technique only requires one of the IQ components to be acquired, it can help to reduce the number of ADCs used in DPD feedback receivers.

The author of this thesis was the main contributor of this paper, performed the experimental work and manuscript writing. The manuscript was reviewed and refined by the co-authors.

Paper E: On the Behavior of the Normalized Mean Square Error in Power Amplifier Linearization

The normalized mean square error (NMSE) is a performance criteria commonly used in power amplifier linearization to quantify the amount of in-band distortion encountered at the output of a power amplifier (PA). This paper presents the derivation of a closed-form expression for the minimum NMSE that could be obtained in a linearization scenario. The derived expression is verified through numerical simulations, and is used to compare the performance of different DPD linearization schemes proposed in the literature.

The author of this thesis was the main contributor of this paper, performed the experimental work and manuscript writing. The manuscript was reviewed and refined by the co-authors.

Acknowledgements

I would like express my gratitude to all the people that made this work possible. First and foremost, I would like to express my gratitude to my main advisor Prof. Thomas Eriksson for giving me the opportunity to pursue a Ph.D. at Chalmers and for all the guidance and support you provided. To my co-advisor Prof. Christian Fager, I am indebted to you for the opportunity you gave me to work with you a long time ago when I was a M.Sc. student and for your support during my Ph.D. studies. To my second co-advisor Dr. Per Landin, thanks for all the many fruitful discussions we had and for all the time you spent reading and discussing my papers, even after you left Chalmers. It was great to work with you.

I would like to thank Prof. Jan Grahn, director of the GigaHertz Centre, for creating such a good research environment and to Prof. Erik Ström, for providing a pleasant working environment here in the Communications and Antenna Systems division. Thanks to my colleagues at the Communication and Antenna Systems division and at the Microwave Electronics Laboratory. Thanks to Dr. Koen Buisman for the support with WebLab measurements. Special thanks goes to Dr. Mustafa Özen for all the fruitful discussions on power amplifiers.

I would also like to thank Agneta Kinnander and Natasha Adler Grønbech for their help in administrative matters, and to Henric Fjellstedt for the computer support in the laboratory.

This work is dedicated to my family and my husband, Mustafa.

This research has been carried out in the GigaHertz Centre in joint projects financed by the Swedish Government Agency for Innovation Systems (VINNOVA), Chalmers University of Technology, Ericsson, Gotmic, Infineon Technologies Austria, National Instruments, Ampleon, Qamcom, RISE, and SAAB.

Bibliography

[1] "5G radio access- capabilities and technologies," Ericsson, White Paper, Apr. 2016. [Online]. Available: https://www.ericsson.com/res/docs/whitepapers/wp-5g.pdf

[2] "5G use cases and requirements," Nokia, Tech. Rep., 2016.

[3] "Big connectivity new future: 2016 sustainability report," China Mobile Limited, Report, Mar. 2017. [Online]. Available: http://www.chinamobileltd.com/en/about/sd/2016/04.pdf

[4] C. L. I, C. Rowell, S. Han, Z. Xu, G. Li, and Z. Pan, "Toward green and soft: a 5G perspective," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 66–73, February 2014.

[5] Energy research & data, "Global energy statistical yearbook 2016," Accessed: 02 July 2017. [Online]. Available: https://yearbook.enerdata.net/#electricity-domestic-consumption-data-by-region.html

[6] United States Environmental Protection Agency, "Greenhouse gas equivalencies," Accessed: 02 July 2017. [Online]. Available: https://www.epa.gov/energy/greenhouse-gas-equivalencies-calculator

[7] "Building zero-emission radio access networks," Nokia, White Paper, 2016. [Online]. Available: https://tools.ext.nokia.com/asset/200830

[8] "C-RAN: The road towards green ran," China Mobile Research Institute, White Paper, Dec. 2013. [Online]. Available: http://labs.chinamobile.com/cran/wp-content/uploads/2014/06/20140613-C-RAN-WP-3.0.pdf

[9] S. N. Roy, "Energy logic: A road map to reducing energy consumption in telecom munications networks," in *Int. Telecommunications Energy Conf.*, Sept 2008, pp. 1–9.

[10] S. Cripps, *RF Power Amplifiers for Wireless Communications.* Artech House, 2006.

[11] A. Mohammadi and F. M. Ghannouchi, *RF Transceiver Design for MIMO Wireless Communications.* Springer Berlin Heidelberg, 2012.

[12] B. Berglun, J. Johansson, and T. Lejon, "High efficiency power amplifiers," Ericsson, White Paper, 2006. [Online]. Available: https://www.ericsson.com/ericsson/corpinfo/publications/review/2006_03/files/1_pa.pdf

[13] D. R. Morgan, Z. Ma, and L. Ding, "Reducing measurement noise effects in digital predistortion of RF power amplifiers," in *IEEE Int. Conf. Commun.*, vol. 4, May 2003, pp. 2436–2439.

[14] P. Lavrador, T. Cunha, P. Cabral, and J. Pedro, "The linearity-efficiency compromise," *IEEE Microw. Mag.*, vol. 11, no. 5, pp. 44–58, Aug. 2010.

[15] P. B. Kenington, "Linearized transmitters: an enabling technology for software defined radio," *IEEE Commun. Mag.*, vol. 40, no. 2, pp. 156–162, Feb. 2002.

[16] D. R. Morgan, Z. Ma, J. Kim, M. G. Z. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.

[17] J. Kim and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electron. Lett.*, vol. 37, no. 23, pp. 1417–1418, Nov. 2001.

[18] K. J. Muhonen, M. Kavehrad, and R. Krishnamoorthy, "Look-up table techniques for adaptive digital predistortion: a development and comparison," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1995–2002, Sep. 2000.

[19] A. Zhu, P. J. Draxler, J. J. Yan, T. Brazil, D. F. Kimball, and P. M. Asbeck, "Open-loop digital predistorter for RF power amplifiers using dynamic deviation reduction-based volterra series," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 7, pp. 1524–1534, July 2008.

[20] O. Hammi and F. M. Ghannouchi, "Power alignment of digital predistorters for power amplifiers linearity optimization," *IEEE Trans. Broadcast.*, vol. 55, no. 1, pp. 109–114, Mar. 2009.

[21] H. Enzinger, K. Freiberger, G. Kubin, and C. Vogel, "A survey of delay and gain correction methods for the indirect learning of digital predistorters," in *IEEE Int. Conf. on Electronics, Circuits Syst.*, Dec. 2016, pp. 285–288.

[22] I. F. Akyildiz, D. M. Gutierrez-Estevez, and E. C. Reyes, "The evolution to 4G cellular systems: LTE-advanced," *Phys. Commun.*, vol. 3, no. 4, pp. 217–244, Dec. 2010.

[23] Yang-Ming Zhu, "Generalized sampling theorem," *IEEE Trans. Circuits Syst. II, Analog and Digital Signal Process.*, vol. 39, no. 8, pp. 587–588, Aug. 1992.

[24] W. A. Frank, "Sampling requirements for Volterra system identification," *IEEE Signal Processing Lett.*, vol. 3, no. 9, pp. 266–268, Sept. 1996.

[25] Y. Liu, J. J. Yan, H. T. Dabag, and P. M. Asbeck, "Novel technique for wideband digital predistortion of power amplifiers with an under-sampling ADC," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 11, pp. 2604–2617, Nov. 2014.

[26] H. Koeppl and P. Singerl, "An efficient scheme for nonlinear modeling and predistortion in mixed-signal systems," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 53, no. 12, pp. 1368–1372, Dec. 2006.

[27] C. Yu, L. Guan, E. Zhu, and A. Zhu, "Band-limited Volterra series-based digital predistortion for wideband RF power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 12, pp. 4198–4208, Dec. 2012.

[28] Q. Zhang, Y. Liu, J. Zhou, S. Jin, W. Chen, and S. Zhang, "A band-divided memory polynomial for wideband digital predistortion with limited bandwidth feedback," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 62, no. 10, pp. 922–926, Oct. 2015.

[29] M. Isaksson, D. Wisell, and D. Ronnow, "A comparative analysis of behavioral models for RF power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 1, pp. 348–359, Jan. 2006.

[30] D. Schreurs, M. O'Droma, A. A. Goacher, and M. Gadringer, *RF Power Amplifier Behavioral Modeling.* Cambridge, 2008.

[31] A. S. Tehrani, H. Cao, S. Afsardoost, T. Eriksson, M. Isaksson, and C. Fager, "A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 6, pp. 1510–1520, June 2010.

[32] F. M. Ghannouchi and O. Hammi, "Behavioral modeling and predistortion," *IEEE Microw. Mag.*, vol. 10, no. 7, pp. 52–64, Dec. 2009.

[33] A. A. M. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Trans. Commun.*, vol. 29, no. 11, pp. 1715–1720, Nov. 1981.

[34] C. Rapp, "Effects of HPA-nonlinearity on a 4-DPSK/OFDM-signal for a digital sound broadcasting system," in *Conf. on Satellite Commun.*, Oct. 1991, pp. 179–184.

[35] K. F. Liang, J. H. Chen, and Y. J. E. Chen, "A quadratic-interpolated lut-based digital predistortion technique for cellular power amplifiers," *IEEE Trans. Circuits Syst. II: Express Briefs*, vol. 61, no. 3, pp. 133–137, March 2014.

[36] M. Schetzen, *The Volterra and Wiener Theorems of Nonlinear Systems.* Wiley, 1980.

[37] A. Zhu and T. J C Pedro, T J Brazil, "Dynamic deviation reduction-based volterra behavioral modeling of RF power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 12, pp. 4323–4332, Dec. 2006.

[38] P. Landin, "Digital baseband modeling and correction of radio frequency power amplifiers," Ph.D. dissertation, KTH Royal Institute of Technology, Stockholm, Sweden, 2012.

[39] S. Haykin, *Digital Communication Systems.* Wiley, 2014.

[40] D. Wisell, J. Jalden, and P. Händel, "Behavioral power amplifier modeling using the LASSO," in *Proc. Conf. Instrum. Meas. Technol.*, May 2008, pp. 1864–1867.

[41] A. Zhu, J. C. Pedro, and T. R. Cunha, "Pruning the volterra series for behavioral modeling of power amplifiers using physical knowledge," *IEEE Trans. Microw. Theory Techn.*, vol. 55, no. 5, pp. 813–821, May 2007.

[42] P. N. Landin, K. Barbé, W. V. Moer, M. Isaksson, and P. Händel, "Two novel memory polynomial models for modeling of rf power amplifiers," *Int. J. Microw. Wireless Technologies*, vol. 7, no. 01, pp. 19–29, 2015.

[43] S. Afsardoost, T. Eriksson, and C. Fager, "Digital predistortion using a vector-switched model," *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 4, pp. 1166–1174, Apr. 2012.

[44] O. Hammi, F. M. Ghannouchi, and B. Vassilakis, "A compact envelope-memory polynomial for RF transmitters modeling with application to baseband and RF-digital predistortion," *IEEE Microw. Wireless Comp. Let.*, vol. 18, no. 5, pp. 359–361, May 2008.

[45] S. M. Kay, *Fundamentals of statistical signal processing, Volume 1: Estimation Theory.* Pearsson, 1993.

[46] G. Lazzarin, S. Pupolin, and A. Sarti, "Nonlinearity compensation in digital radio systems," *IEEE Trans.Commun.*, vol. 42, no. 234, pp. 988–999, Feb. 1994.

[47] C. Eun and E. J. Powers, "A new Volterra predistorter based on the indirect learning architecture," *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 223–227, Jan. 1997.

[48] M. Schetzen, "Theory of pth-order inverses of nonlinear systems," *IEEE Trans. Circuits Syst.*, vol. 23, no. 5, pp. 285–291, May 1976.

[49] P. B. Kenington, "Tiny digital predistortion receiver integrates RF, filter and ADC," *LT J. of Analog Innovation*, 2010.

[50] P. N. Landin, S. Gustafsson, C. Fager, and T. Eriksson, "Weblab: A web-based setup for PA digital predistortion and characterization [application notes]," *IEEE Microw. Mag.*, vol. 16, no. 1, pp. 138–140, Feb 2015.

[51] P. N. Landin, M. Isaksson, and P. Händel, "Parameter extraction and performance evaluation method for increased performance in RF power amplifier behavioral modeling," *Int. Journal of RF and Microw. Computer-Aided Engineering*, vol. 20, no. 2, pp. 200–208, 2010. [Online]. Available: http://dx.doi.org/10.1002/mmce.20422

[52] F. Gregorio, J. Cousseau, S. Werner, T. Riihonen, and R. Wichman, "EVM analysis for broadband OFDM direct-conversion transmitters," *IEEE Trans. Veh. Technol.*, vol. 62, no. 7, pp. 3443–3451, Sept. 2013.

[53] D. A. Bristow, M. Tharayil, and A. G. Alleyne, "A survey of iterative learning control," *IEEE Control Syst. Mag.*, vol. 26, no. 3, pp. 96–114, June 2006.

[54] H. S. Ahn, K. L. Moore, and Y. Chen, *Iterative Learning Control: Robustness and Monotonic Convergence for Interval Systems.* Springer, 2007.

[55] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 159–165, Jan. 2004.

[56] ——, "Memory polynomial predistorter based on the indirect learning architecture," in *IEEE Global Telecomm. Conf. GLOBECOM*, vol. 1, Nov 2002, pp. 967–971.

[57] M. Isaksson and D. Rönnow, "A parameter-reduced volterra model for dynamic rf power amplifier modeling based on orthonormal basis functions," in *In press.* In Press, 2007.

[58] E. Zenteno, S. Amin, M. Isaksson, D. Rönnow, and P. Händel, "Combating the dimensionality of nonlinear MIMO amplifier predistortion by basis pursuit," in *European Microw. Conf.*, Oct 2014, pp. 833–836.

[59] J. Reina-Tosina, M. Allegue-Martinez, M. Madero-Ayora, C. Crespo-Cadenas, and S. Cruces, "Digital predistortion based on a compressed-sensing approach," in *European Microw. Conf.*, Oct. 2013, pp. 408–411.

[60] J. Reina-Tosina, M. Allegue-Martinez, C. Crespo-Cadenas, C. Yu, and S. Cruces, "Behavioral modeling and predistortion of power amplifiers under sparsity hypothesis," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 2, pp. 745–753, Feb. 2015.

[61] A. Abdelhafiz, A. Kwan, O. Hammi, and F. M. Ghannouchi, "Digital predistortion of LTE-A power amplifiers using compressed-sampling-based unstructured pruning of volterra series," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 11, pp. 2583–2593, Nov. 2014.

[62] D. Psaltis, A. Sideris, and A. A. Yamamura, "A multilayer neural network controller," *IEEE Control Syst. Mag.*, 1988.

[63] P. N. Landin, A. E. Mayer, and T. Eriksson, "MILA - a noise mitigation technique for RF power amplifier linearization," in *Int. Multi-Conf. Syst., Signals Devices*, Feb. 2014, pp. 1–4.

[64] S. Amin, E. Zenteno, P. N. Landin, D. Rönnow, M. Isaksson, and P. H andel, "Noise impact on the identification of digital predistorter parameters in the indirect learning architecture," in *Swedish Commun. Technologies Workshop Dig.*, Oct. 2012, pp. 36–39.

[65] D. Zhou and V. E. DeBrunner, "Novel adaptive nonlinear predistorters based on the direct learning algorithm," *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 120–133, Jan. 2007.

[66] H. Passo and A. Mammela, "Comparison of direct learning and indirect learning predistortion architectures," in *IEEE Int. Symp. Wireless Commun. Syst.*, Oct. 2008, pp. 309–313.

[67] M. A. Hussein, V. A. Bohara, and O. Venard, "On the system level convergence of ILA and DLA for digital predistortion," in *Int. Symp. on Wireless Commun. Syst.*, Aug. 2012, pp. 870–874.

[68] Y. H. Lim, Y.-S. Cho, I. W. Cha, and D.-H. Youn, "An adaptive nonlinear prefilter for compensation of distortion in nonlinear systems," *IEEE Trans. Signal Process.*, vol. 46, no. 6, pp. 1726–1730, June 1998.

[69] H.-W. Kang, Y.-S. Cho, and D.-H. Youn, "On compensating nonlinear distortions of an OFDM system using an efficient adaptive predistorter," *IEEE Trans. Commun.*, vol. 47, no. 4, pp. 522–526, Apr. 1999.

[70] R. N. Braithwaite, "Reducing estimator biases due to equalization errors in adaptive digital predistortion systems for RF power amplifiers," in *IEEE MTT-S Int. Microwave Symp. Dig.*, June 2012, pp. 1–3.

[71] ——, "A comparison of indirect learning and closed loop estimators used in digital predistortion of power amplifiers," in *2015 IEEE MTT-S International Microwave Symposium*, May 2015, pp. 1–4.

[72] J. Chani-Cahuana, C. Fager, and T. Eriksson, "A new variant of the indirect learning architecture for the linearization of power amplifiers," in *Proc. European Microw. Integrated Circuits Conf.*, Sept. 2015, pp. 444–447.

[73] C. H. A. Tavares, J. C. M. Filho, C. M. Panazio, and T. Abrão, "Input back-off optimization in OFDM systems under ideal pre-distorters," *IEEE Wireless Commun. Lett.*, vol. 5, no. 5, pp. 464–467, Oct. 2016.

[74] W. T. Lin and D. C. Chang, "Adaptive carrier synchronization using decision-aided kalman filtering algorithms," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 4, pp. 1260–1267, Nov. 2007.